CRITICAL FUNCTIONS AND BLOW-UP ASYMPTOTICS FOR THE FRACTIONAL BREZIS–NIRENBERG PROBLEM IN LOW DIMENSION

NICOLA DE NITTI AND TOBIAS KÖNIG

ABSTRACT. For $s \in (0,1)$ and a bounded open set $\Omega \subset \mathbb{R}^N$ with $N > 2s$, we study the fractional Brezis–Nirenberg type minimization problem of finding

$$S(a) := \inf \frac{\int_{\mathbb{R}^N} |(-\Delta)^{s/2} u|^2 + \int_\Omega a u^2}{\left(\int_\Omega u^{\frac{2N}{N-2s}} \right)^{\frac{N-2s}{N}}},$$

where the infimum is taken over all functions $u \in H^s(\mathbb{R}^N)$ that vanish outside Ω. The function a is assumed to be critical in the sense of Hebey and Vaugon. For low dimensions $N \in (2s, 4s)$, we prove that the Robin function ϕ_a satisfies $\inf_{x \in \Omega} \phi_a(x) = 0$, which extends a result obtained by Druet for $s = 1$. In dimensions $N \in (8s/3, 4s)$, we then study the asymptotics of the fractional Brezis–Nirenberg energy $S(a + \varepsilon V)$ for some $V \in L^\infty(\Omega)$ as $\varepsilon \to 0+$. We give a precise description of the blow-up profile of (almost) minimizing sequences and characterize the concentration speed and the location of concentration points.

1. INTRODUCTION AND MAIN RESULTS

Let $N \in \mathbb{N}$ and $0 < 2s < N$ for some $s \in (0,1)$, and let $\Omega \subset \mathbb{R}^N$ be a bounded open set. The goal of the present paper is to analyze the variational problem of minimizing, for a given $a \in C(\overline{\Omega})$, the quotient functional

$$\mathcal{S}_a[u] := \frac{\int_{\mathbb{R}^N} |(-\Delta)^{s/2} u|^2 \, \mathrm{d}y + \int_\Omega a(y) u(y)^2 \, \mathrm{d}y}{\|u\|^2_{L^{\frac{2N}{N-2s}}(\Omega)}} \tag{1.1}$$

over functions in the space

$$\widetilde{H}^s(\Omega) := \left\{ u \in H^s(\mathbb{R}^N) : u \equiv 0 \quad \text{on } \mathbb{R}^N \setminus \Omega \right\}, \tag{1.2}$$

where $u \in H^s(\mathbb{R}^N)$ iff

$$\|u\|_{L^2(\mathbb{R}^N)} + \left(\int_{\mathbb{R}^N} |(-\Delta)^{s/2} u|^2 \, \mathrm{d}y \right)^{1/2} < \infty, \tag{1.3}$$

and the fractional Laplacian operator $(-\Delta)^s u$ is defined for any $u \in H^s(\mathbb{R}^N)$ through the Fourier representation

$$(-\Delta)^s u = \mathcal{F}^{-1}(|\xi|^{2s} \mathcal{F} u). \tag{1.4}$$

We also recall the singular integral representation of the fractional Laplacian (see [11, 25]):

$$(-\Delta)^s u(x) := C_{N,s} P.V. \int_{\mathbb{R}^N} \frac{u(x) - u(y)}{|x - y|^{N+2s}} \, \mathrm{d}y, \tag{1.5}$$

where

$$C_{N,s} := \frac{s 2^{2s} \Gamma(\frac{N+2s}{2})}{\pi^{N/2} \Gamma(1-s)}. \tag{1.6}$$

The associated infimum,

$$S(a) := \inf \left\{ \mathcal{S}_a[u] \ : \ u \in \widetilde{H}^s(\Omega) \right\}, \tag{1.7}$$

is to be compared with the number $S := S_{N,s} := S(0)$, which is equal to the best constant in the fractional Sobolev embedding

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2} u|^2 \, \mathrm{d}y \geq S \|u\|^2_{L^{\frac{2N}{N-2s}}(\mathbb{R}^N)}, \tag{1.8}$$

given by

$$S_{N,s} := 2^{2s} \pi^s \frac{\Gamma(\frac{N+2s}{2})}{\Gamma(\frac{N-2s}{2})} \left(\frac{\Gamma(N/2)}{\Gamma(N)} \right)^{2s/N}. \tag{1.9}$$

We note that the embedding $\widetilde{H}^s(\Omega) \hookrightarrow L^{p+1}(\Omega)$ and the associated best constant are in fact independent of Ω and equal to the best full-space Sobolev constant $S_{N,s}$ (see [19]). In the classical case $s = 1$, problem (1.7) has been first studied in the famous paper [9] by Brezis and Nirenberg, who were interested in obtaining positive solutions to the associated elliptic equation. One of the main findings in that paper is that the behavior of (1.7) depends on the space dimension N in a rather striking way. Indeed, when $N \geq 4$, then $S(a) < S$ if and only if $a(x) < 0$ for some $x \in \Omega$. On the other hand, when $N = 3$, then $S(a) = S$ whenever $\|a\|_\infty$ is small enough, leaving open the question of characterizing the cases $S(a) < S$ in terms of a. In [20], Druet proved that, for $N = 3$, the following equivalence holds:

$$S(a) < S \qquad \Longleftrightarrow \qquad \phi_a(x) < 0 \quad \text{for some } x \in \Omega, \tag{1.10}$$

where $\phi_a(x)$ denotes the Robin function associated to a (see (1.11) below). This answered positively a conjecture previously formulated by Brezis in [8].

For a fractional power $s \in (0, 1)$ and assuming $a = -\lambda$ for some constant $\lambda > 0$, Brezis–Nirenberg type results have been obtained by Servadei and Valdinoci:

  (i) In [38], they proved that when $N \geq 4s$, then $S(-\lambda) < S$ whenever $\lambda > 0$;
  (ii) In [36], they proved that for $2s < N < 4s$ and prove that there is $\lambda_s \in (0, \lambda_{1,s})$ (where $\lambda_{1,s}$ is the first Dirichlet eigenvalue of $(-\Delta)^s$) such that for every $\lambda \in (\lambda_s, \lambda_{1,s})$, one has $S(-\lambda) < S$.

In this paper, we shall exclusively be concerned with the low-dimensional range $2s < N < 4s$. This is the natural replacement of the classical case $N = 3$, $s = 1$, as indicated by the results above. One may also notice that when $2s < N$, the Green's function for $(-\Delta)^s$ on $\mathbb{R}^N$ behaves like $G(x, y) \sim |x-y|^{-N+2s}$ near the diagonal and thus fails to be in $L^2_{\text{loc}}(\mathbb{R}^N)$ precisely if $N \leq 4s$, compare [29].

A central notion to what follows is that of a critical function a, which was introduced by Hebey and Vaugon in [28] for for $s = 1$ and readily generalizes to the fractional situation. Indeed, the following definition is naturally suggested by the behavior of $S(a)$ just described.

Definition 1.1 (Critical function). Let $a \in C(\overline{\Omega})$. We say that a is *critical* if $S(a) = S$ and $S(\tilde{a}) < S(a)$ for every $\tilde{a} \in C(\overline{\Omega})$ with $\tilde{a} \leq a$ and $\tilde{a} \not\equiv a$.

When $N \geq 4s$, the result of [38] implies that the only critical potential is $a \equiv 0$. For this case, or more generally for $N > 2s$ with $a \equiv 0$, the recent literature is rather rich in refined results going beyond [38]. Notably, in [15] and [14], the authors prove the fractional counterpart of some conjectures by Brezis and Peletier [10] concerning the blow-up asymptotics of minimizers to the problem $S(-\varepsilon)$ and a related problem with subcritical exponent $p - \varepsilon$ as $\varepsilon \to 0$. In the classical case $s = 1$, these results are due to Han [27] and Rey [33, 34]. Corresponding existence results, also for non-minimizing multi-bubble solutions, are also given in [15, 14], as well as in [17, 26].

In contrast to this, in the more challenging setting of dimension $2s < N < 4s$, critical functions can have all possible shapes and are necessarily non-zero, compare [20] and Corollary 1.3 below. In this setting, and notably in the presence of a critical function, results of Han–Rey type as just discussed

are much more scarce in the literature. Even in the local case $s = 1$ and $N = 3$, the conjecture of Brezis and Peletier (see [10, Conjecture 3.(ii)]) which involves a (constant) critical function has only been proved recently in [23]. For the fractional case $2s < N < 4s$, we are not aware of any results going beyond [36]. The purpose of the present paper is to start filling this gap.

1.1. **Main results.** For all of our results, a crucial role is played by the Green's function of $(-\Delta)^s + a$, which we introduce now. For a function $a \in C(\overline{\Omega})$ such that $(-\Delta)^s + a$ is coercive, i.e.

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2} v|^2 \, \mathrm{d}y + \int_\Omega av^2 \, \mathrm{d}y \geq c \int_{\mathbb{R}^N} |(-\Delta)^{s/2} v|^2 \, \mathrm{d}y$$

for some $c > 0$, define $G_a : \Omega \times \mathbb{R}^N \to \mathbb{R}$ as the unique function such that for every fixed $x \in \Omega$

$$\begin{cases} ((-\Delta)^s + a)G_a(x, \cdot) = \gamma_{N,s}\delta_x & \text{in } \Omega, \\ G_a(x, \cdot) = 0 & \text{on } \mathbb{R}^N \setminus \Omega. \end{cases}$$

Here, we set $\gamma_{N,s} = \frac{2^{2s}\pi^{N/2}\Gamma(s)}{\Gamma(\frac{N-2s}{2})}$, so that $(-\Delta)^s |y|^{-N+2s} = \gamma_{N,s}\delta_0$ on $\mathbb{R}^N$. Thus this choice of $\gamma_{N,s}$ ensures that we can write G_a as a sum of its singular part and its regular part $H_a(x, y)$ as follows:

$$G_a(x, y) = \frac{1}{|x - y|^{N-2s}} - H_a(x, y).$$

The function H_a is continuous up to the diagonal, see e.g. Lemma A.3. Therefore, we may define the *Robin function*

$$\phi_a(x) := H_a(x, x), \qquad x \in \Omega. \tag{1.11}$$

We prove several properties of the Green's functions G_a in Appendix A.

Our first main result is the following extension of Druet's theorem from [20] to the fractional case.

Theorem 1.2 (Characterization of criticality). *Let $2s < N < 4s$ and let $a \in C(\overline{\Omega})$ be such that $(-\Delta)^s + a$ is coercive. The following properties are equivalent.*

(i) There is $x \in \Omega$ such that $\phi_a(x) < 0$.
(ii) $S(a) < S$.
(iii) $S(a)$ is achieved by some function $u \in \widetilde{H}^s(\Omega)$.

As an immediate corollary, we can characterize critical functions in terms of their Robin function.

Corollary 1.3. *Let a be critical. Then $\inf_{x \in \Omega} \phi_a(x) = 0$.*

The implications $(i) \Rightarrow (ii)$ and $(ii) \Rightarrow (iii)$ in Theorem 1.2 are well-known: indeed, $(i) \Rightarrow (ii)$ easily follows by the proper choice of test functions thanks to Theorem 2.1 below; the implication $(ii) \Rightarrow (iii)$ is the fractional version of the seminal observation in [9] (see [38, Theorem 2]).

Our proof of $(iii) \Rightarrow (ii)$ is the content of Proposition 3.1 below and follows [20, Step 1]. The most involved proof is that implication is $(ii) \Rightarrow (i)$, which we give in Section 4. We adapt the strategy developed by Esposito in [21], who gave an alternative proof of that implication. His approach is based on an expansion of the energy functional $\mathcal{S}_{a-\varepsilon}[u_\varepsilon]$ as $\varepsilon \to 0$, where a is critical as in Definition 1.1 and u_ε is a minimizer of $S(a - \varepsilon)$.

In fact, by using the techniques applied in the recent work [24] for $s = 1$, we are even able to push this expansion of $\mathcal{S}_{a-\varepsilon}[u_\varepsilon]$ further by one order of ε and derive precise asymptotics of the energy $S(a - \varepsilon)$ and of the sequence (u_ε).

To give a precise statement of our results, let us fix some more assumptions and notations. We denote the zero set of the Robin function ϕ_a by

$$\mathcal{N}_a := \{x \in \Omega : \phi_a(x) = 0\}.$$

It follows from Theorem 1.2 that $\inf_\Omega \phi_a(x) = 0$ if and only if a is critical. In particular, $\mathcal{N}_a$ is not empty if a is critical.

We will consider perturbations of a of the form $a + \varepsilon V$, with non-constant $V \in L^\infty(\Omega)$. For such V, following [24], we let

$$Q_V(x) := \int_\Omega V(y) G_a(x, y)^2 \, \mathrm{d}y$$

and

$$\mathcal{N}_a(V) := \{x \in \mathcal{N}_a \,:\, Q_V(x) < 0\}.$$

Finally, we shall assume that Ω has C^2 boundary and that

$$a \in C(\bar{\Omega}) \cap C^1(\Omega), \ a(x) < 0 \quad \text{for all } x \in \mathcal{N}_a; \tag{1.12}$$

By Corollary 2.2, we have a priori that $a(x) \le 0$ on $\mathcal{N}_a$. Therefore assumption (1.12) is not severe.

We point out that with our methods we are able to prove the following theorems only for the restricted dimensional range $\frac{8}{3}s < N < 4s$, which enters in Section 5. We discuss this assumption in some more detail after the statement of Theorem 1.6 below.

The following theorem describes the asymptotics of the perturbed minimal energy $S(a + \varepsilon V)$ as $\varepsilon \to 0+$. It shows in particular the non-obvious fact that the condition $\mathcal{N}_a(V) \ne \emptyset$ is sufficient to have $S(a + \varepsilon V) < S$.

Theorem 1.4 (Energy asymptotics)**.** *Let $\frac{8}{3}s < N < 4s$. Let us assume that $\mathcal{N}_a(V) \ne \emptyset$. Then $S(a + \varepsilon V) < S$ for all $\varepsilon > 0$ and*

$$\lim_{\varepsilon \to 0^+} \frac{S(a + \varepsilon V) - S}{\varepsilon^{\frac{2s}{4s-N}}} = \sigma_{N,s} \sup_{x \in \mathcal{N}_a(V)} \frac{|Q_V(x)|^{\frac{2s}{4s-N}}}{|a(x)|^{\frac{N-2s}{4s-N}}},$$

where $\sigma_{N,s} > 0$ is a dimensional constant given explicitly by

$$\sigma_{N,s} = A_{N,s}^{-\frac{N-2s}{N}} (\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s})^{-\frac{N-2s}{4s-N}} \left(\frac{N-2s}{2s}\right)^{\frac{2s}{4s-N}} \frac{4s-N}{N-2s}.$$

The constants $A_{N,s}$, $\alpha_{N,s}$, $c_{N,s}$, $d_{N,s}$ and $b_{N,s}$ are given explicitly in Lemma B.5 below.

On the other hand, when $\mathcal{N}_a(V) = \emptyset$, the next theorem shows that the asymptotics become trivial provided $Q_V > 0$ on $\mathcal{N}_a$. Only in the case when $\min_{\mathcal{N}_a} Q_V = 0$ we do not obtain the precise leading term of $S(a + \varepsilon V) - S$.

Theorem 1.5 (Energy asymptotics, degenerate case)**.** *Let $\frac{8}{3}s < N < 4s$. Let us assume that $\mathcal{N}_a(V) = \emptyset$. Then $S(a+\varepsilon V) = S + o(\varepsilon^2)$ as $\varepsilon \to 0^+$. If, in addition, $Q_V(x) > 0$ for all $x \in \mathcal{N}_a$ then $S(a+\varepsilon V) = S$ for sufficiently small $\varepsilon > 0$.*

For a potential V such that $\mathcal{N}_a(V) \ne \emptyset$, and thus $S(a + \varepsilon V) < S$ by Theorem 1.4, a minimizer u_ε of $S(a + \varepsilon V)$ exists by Theorem 1.2. We now turn to studying the asymptotic behavior of the sequence (u_ε). In fact, since our methods are purely variational, we do not need to require that the u_ε satisfy a corresponding equation and we can equally well treat a sequence of almost minimizers in the sense of (1.17) below.

Since the functional $\mathcal{S}_a$ is merely a perturbation of the standard Sobolev quotient functional, it is not surprising that to leading order, the sequence u_ε approaches the family of functions

$$U_{x,\lambda}(y) = \left(\frac{\lambda}{1 + \lambda^2 |x - y|^2}\right)^{\frac{N-2s}{2}}, \qquad x \in \mathbb{R}^N, \quad \lambda > 0. \tag{1.13}$$

The $U_{x,\lambda}$ are precisely the optimizers of the fractional Sobolev inequality on $\mathbb{R}^N$

$$\|(-\Delta)^{s/2}u\|_{L^2(\mathbb{R}^N)}^2 \geq S_{N,s}\|u\|_{L^{\frac{2N}{N-2s}}(\mathbb{R}^N)}^2. \tag{1.14}$$

and satisfy the equation

$$(-\Delta)^s U_{x,\lambda}(y) = c_{N,s}U_{x,\lambda}(y)^{\frac{N+2s}{N-2s}}, \tag{1.15}$$

with $c_{N,s} > 0$ given explicitly in Lemma B.5.

Since we are working on the bounded set Ω, the first refinement of the approximation consists in 'projecting' the functions $U_{x,\lambda}$ to $\widetilde{H}^s(\Omega)$. That is, we consider the unique function $PU_{x,\lambda} \in \widetilde{H}^s(\Omega)$ satisfying

$$\begin{cases} (-\Delta)^s PU_{x,\lambda} = (-\Delta)^s U_{x,\lambda} & \text{in } \Omega, \\ PU_{x,\lambda} = 0 & \text{on } \mathbb{R}^N \setminus \Omega. \end{cases} \tag{1.16}$$

in the weak sense, that is,

$$\int_{\mathbb{R}^N} (-\Delta)^{s/2}PU_{x,\lambda}(-\Delta)^{s/2}\eta\,\mathrm{d}y = \int_{\mathbb{R}^N} (-\Delta)^s U_{x,\lambda}\eta\,\mathrm{d}y = c_{N,s}\int_\Omega U_{x,\lambda}^{\frac{N+2s}{N-2s}}\eta\,\mathrm{d}y$$

for every $\eta \in \widetilde{H}^s(\Omega)$.

Finally, we introduce the space

$$T_{x,\lambda} = \operatorname{span}\left\{PU_{x,\lambda}, \partial_\lambda PU_{x,\lambda}, \{\partial_{x_i}PU_{x,\lambda}\}_{i=1}^N\right\} \subset \widetilde{H}^s(\Omega)$$

and denote by $T_{x,\lambda}^\perp \subset \widetilde{H}^s(\Omega)$ its orthogonal complement in $\widetilde{H}^s(\Omega)$ with respect to the scalar product $(u,v) := \int_{\mathbb{R}^N} (-\Delta)^{s/2}u(-\Delta)^{s/2}v\,\mathrm{d}y$. Moreover, let us denote by $\Pi_{x,\lambda}$ and $\Pi_{x,\lambda}^\perp$ the projections onto $T_{x,\lambda}$ and $T_{x,\lambda}^\perp$ respectively.

Then we have the following result.

Theorem 1.6 (Concentration of almost-minimizers). *Let $\frac{8}{3}s < N < 4s$. Suppose that $(u_\varepsilon) \subset \widetilde{H}^s(\Omega)$ is a sequence such that*

$$\lim_{\varepsilon\to 0}\frac{\mathcal{S}_{a+\varepsilon V}[u_\varepsilon] - S(a+\varepsilon V)}{S - S(a+\varepsilon V)} = 0 \quad \text{and} \quad \int_\Omega u_\varepsilon^p\,\mathrm{d}y = \int_{\mathbb{R}^N} U_{0,1}^p\,\mathrm{d}y. \tag{1.17}$$

Then there exist sequences $(x_\varepsilon) \subset \Omega$, $(\lambda_\varepsilon) \subset (0,\infty)$, $(w_\varepsilon) \subset T_{x_\varepsilon,\lambda_\varepsilon}^\perp$, and $(\alpha_\varepsilon) \subset \mathbb{R}$ such that, up to extraction of a subsequence,

$$u_\varepsilon = \alpha_\varepsilon\left(PU_{x_\varepsilon,\lambda_\varepsilon} + \lambda^{-\frac{N-2s}{2}}\Pi_{x_\varepsilon,\lambda_\varepsilon}^\perp\left(H_0(x_\varepsilon,\cdot) - H_a(x_\varepsilon,\cdot)\right) + r_\varepsilon\right) \tag{1.18}$$

Moreover, as $\varepsilon \to 0$, we have

$$x_\varepsilon \to x_0 \quad \text{for some } x_0 \text{ such that} \quad \frac{|Q_V(x_0)|^{\frac{2s}{4s-N}}}{|a(x_0)|^{\frac{N-2s}{4s-N}}} = \sup_{y\in\mathcal{N}_a(V)}\frac{|Q_V(y)|^{\frac{2s}{4s-N}}}{|a(y)|^{\frac{N-2s}{4s-N}}},$$

$$\phi_a(x_\varepsilon) = o(\varepsilon),$$

$$\lim_{\varepsilon\to 0}\varepsilon^{\frac{1}{4s-N}}\lambda_\varepsilon = \left(\frac{2s(\alpha_{N,s} + c_{N,s}d_{N,s}b_{N,s})|a(x_0)|}{(N-2s)|Q_V(x_0)|}\right)^{\frac{1}{4s-N}},$$

$$\alpha_\varepsilon = \xi + \mathcal{O}(\varepsilon^{\frac{N-2s}{4s-N}}) \quad \text{for some } \xi \in \{\pm 1\}.$$

Finally, $r_\varepsilon \in T_{x_\varepsilon,\lambda_\varepsilon}^\perp$ and $\|(-\Delta)^{\frac{s}{2}}r\|_{L^2(\mathbb{R}^N)}^2 = o(\varepsilon^{\frac{2s}{4s-N}})$.

The constants $\alpha_{N,s}$, $c_{N,s}$, $d_{N,s}$ and $b_{N,s}$ are given explicitly in Lemma B.5.

Theorem 1.6 should be seen as the low-dimensional counterpart of [15, Theorems 1.1 and 1.2], which concerns $N > 4s$. The decisive additional complication to be overcome in our case is the presence of a non-zero critical function a. More concretely, the coefficient $\phi_a(x)$ of the subleading term of the energy expansion vanishes due to criticality of a (compare Theorem 2.1 and Lemma 5.5). As a consequence, it is only after further refining the expansion that we are able to conclude the desired information about the concentration behavior of the sequence u_ε.

In the same vein, the energy expansions from Theorem 1.4 are harder to obtain than their analogues in higher dimensions $N \geq 4s$. Indeed, for $N > 4s$ we have

$$S(\varepsilon V) = S_{N,s} - \tilde{c}_{N,s} \sup_{\{x \in \Omega \, : \, V(x) < 0\}} \phi_0(x)^{-\frac{2s}{N-4s}} |V(x)|^{\frac{N-2s}{N-4s}} \varepsilon^{\frac{N-2s}{N-4s}} + o(\varepsilon^{\frac{N-2s}{N-4s}}), \qquad (1.19)$$

where $\tilde{c}_{N,s} > 0$ is some dimensional constant. In this case, a sharp upper bound on $S(\varepsilon V)$ can already be derived from testing $\mathcal{S}_{\varepsilon V}$ against the family of functions $PU_{x,\lambda}$. In contrast, for $2s < N < 4s$ this family needs to be modified by a lower order term in order give the sharp upper bound for Theorem 1.4 (see (2.1) and Theorem 2.1 below). For details of the computations in case $N \geq 4s$, we refer to the forthcoming work [18]. It is noteworthy that the auxiliary minimization problem giving the subleading coefficient in (1.19) is local in V in the sense that it only involves the pointwise value $V(x)$, whereas that of Theorem 1.4 contains the non-local quantity Q_V.

Let us now describe in more detail the approach we use in the proofs of Theorems 1.4, 1.5 and 1.6, which are in fact intimately linked. Firstly, the family of functions $\psi_{x,\lambda}$ defined in (2.1) below yields an upper bound for $S(a + \varepsilon V)$, which will turn out to be sharp. The strategy we use to prove the corresponding lower bound on $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$, for a sequence (u_ε) as in (1.17), can be traced back at least to work of Rey [33, 34] and Bahri–Coron [3] on the classical Brezis–Nirenberg problem for $s = 1$; it was adapted to treat problems with a critical potential a when $s = 1$, $N = 3$ in [21] and, more recently, in [24, 23]. This strategy features two characteristic steps, namely (i) supplementing the initial asymptotic expansion $u_\varepsilon = \alpha_\varepsilon(PU_{x_\varepsilon,\lambda_\varepsilon} + w_\varepsilon)$, obtained by a concentration-compactness argument, by the orthogonality condition $w_\varepsilon \in T^\perp_{x_\varepsilon,\lambda_\varepsilon}$ and (ii) using a certain coercivity inequality, valid for functions in $T^\perp_{x_\varepsilon,\lambda_\varepsilon}$, to improve the bound on the remainder w_ε. The basic instance of this strategy is carried out in Section 3. Indeed, after performing steps (i) and (ii), in Proposition 3.6 below we are able to exclude concentration near $\partial\Omega$ and obtain a quantitative bound on $w_\varepsilon = \alpha_\varepsilon^{-1} u_\varepsilon - PU_{x_\varepsilon,\lambda_\varepsilon}$. As in [33] and [22], this information is enough to arrive at (1.19) and similar conclusions when $N > 4s$; see the forthcoming paper [18] for details.

On the other hand, when $2s < N < 4s$, the bound that Proposition 3.6 provides for the modified difference $u_\varepsilon - \psi_{x_\varepsilon,\lambda_\varepsilon}$ is still insufficient. For $s = 1$, it was however observed in [24] that one can refine the expansion of u_ε by reiterating steps (i) and (ii). Here, we carry out their strategy in a streamlined version (compare Remark 5.1) and for fractional $s \in (0,1)$. That is, one writes $w_\varepsilon = \psi_{x_\varepsilon,\lambda_\varepsilon} - PU_{x_\varepsilon,\lambda_\varepsilon} + q_\varepsilon$, decomposes $q_\varepsilon = t_\varepsilon + r_\varepsilon$ with $r_\varepsilon \in T^\perp_{x_\varepsilon,\lambda_\varepsilon}$ and applies the coercivity inequality a second time. We are able to conclude as long as the technical condition $8s/3 < N$ is met (which is equivalent to $\lambda^{-3N+6s} = o(\lambda^{-2s})$). Indeed, in that case the leading contributions of t_ε to the energy, which enter to orders λ^{-N+2s} and λ^{-2N+4s}, cancel precisely; see Lemma 5.8. If $8s/3 \leq N$, a plethora of additional terms in t_ε, which contribute to orders $\lambda^{-k(N-2s)}$ with $3 \leq k \leq \frac{2s}{N-2s}$, will become relevant, and we were not able to treat those in a systemized way. It is natural to expect that the cancellation phenomenon that occurs for $k = 1, 2$ still persists for $k \geq 3$. This would allow to prove Theorems 1.4, 1.5, and 1.6 for general $N > 2s$. For further details of the argument and the difficulties just discussed, we refer to Section 5.

As far as we know, the role of the threshold configurations given by $k(N - 2s) = 2s$ for $k \geq 1$ in the fractional Brezis–Nirenberg problem (1.7) has only been investigated in the literature for $k = 1$ corresponding to $N = 4s$, below which the problem is known to behave differently by the results quoted above. It would be exciting to exhibit some similar, possibly refined, qualitative change in the behavior of (1.7) at one or each of the following thresholds $N = 3s$, $N = 8s/3$, $N = 10s/4$, etc.

To conclude this introduction, let us mention that several works in the literature (see [41, 4, 6]) treat the problem corresponding to (1.7) for a different notion of Dirichlet boundary conditions for $(-\Delta)^s$ on Ω, namely the *spectral* fractional Laplacian, defined by classical spectral theory using the $L^2(\Omega)$-ONB of Dirichlet eigenfunctions for $-\Delta$. In contrast to this, the notion of $(-\Delta)^s$ we use in this paper, as defined in (1.4) or (1.5) on $\widetilde{H}^s(\Omega)$ given by (1.2), usually goes in the literature by the name of *restricted* fractional Laplacian. A nice discussion of these two operators, as well as a method of unified treatment for both, can be found in [17] (see also [37]).

Our method of proof just described is rather different from most of the previous contributions to the fractional Brezis–Nirenberg problem. Namely, we do not need to pass through the extension formulation for $(-\Delta)^s$ due to either [13] for the restricted or to [12, 40] for the spectral fractional Laplacian. On the other hand, using the properties of $PU_{x,\lambda}$ (as given in Lemma A.2) allows us to avoid lengthy calculations with singular integrals, appearing e.g. in [38], while at the same time optimizing the cutoff procedure with respect to [38].

1.2. **Notation.** We will often abbreviate the fractional critical Sobolev exponent by $p := \frac{2N}{N-2s}$. For any $q \geq 1$, we abbreviate $\| \cdot \|_q := \| \cdot \|_{L^q(\mathbb{R}^N)}$. When $q = 2$, we sometimes write $\| \cdot \| := \| \cdot \|_2$.

Unless stated otherwise, we shall always assume $s \in (0,1)$ and $N \in (2s, 4s)$.

For $x \in \Omega$, we use the shorthand $d(x) = \text{dist}(x, \partial\Omega)$.

For a set M and functions $f, g : M \to \mathbb{R}_+$, we shall write $f(m) \lesssim g(m)$ if there exists a constant $C > 0$, independent of m, such that $f(m) \leq Cg(m)$ for all $m \in M$, and accordingly for $\gtrsim$. If $f \lesssim g$ and $g \lesssim f$, we write $f \sim g$.

The various constants appearing throughout the paper and their numerical values are collected in Lemma B.5 in the appendix.

2. Proof of the upper bound

The following theorem gives the asymptotics of $\mathcal{S}_{a+\varepsilon V}[\psi_{x,\lambda}]$, for the test function

$$\psi_{x,\lambda}(y) := PU_{x,\lambda}(y) - \lambda^{-\frac{N-2s}{2}}(H_a(x,y) - H_0(x,y)), \tag{2.1}$$

as $\lambda \to \infty$.

Theorem 2.1 (Expansion of $\mathcal{S}_{a+\varepsilon V}[\psi_{x,\lambda}]$). *As $\lambda \to \infty$, uniformly for x in compact subsets of Ω and for $\varepsilon \geq 0$,*

$$\|(-\Delta)^{s/2}\psi_{x,\lambda}\|_2^2 + \int_\Omega (a + \varepsilon V)\psi_{x,\lambda}^2 \, dy$$
$$= c_{N,s}A_{N,s} - c_{N,s}a_{N,s}\phi_a(x)\lambda^{-N+2s} \tag{2.2}$$
$$+ (c_{N,s}d_{N,s}b_{N,s} - \alpha_{N,s})a(x)\lambda^{-2s} + \varepsilon\lambda^{-N+2s}Q_V(x) + o(\lambda^{-2s}) + o(\varepsilon\lambda^{-N+2s})$$

and

$$\int_\Omega \psi_{x,\lambda}^p \, dy = A_{N,s} - pa_{N,s}\phi_a(x)\lambda^{-N+2s} + \mathcal{T}_1(\phi_a(x), \lambda) + pd_{N,s}b_{N,s}a(x)\lambda^{-2s} + o(\lambda^{-2s}). \tag{2.3}$$

In particular,

$$\mathcal{S}_{a+\varepsilon V}[\psi_{x,\lambda}] = S + A_{N,s}^{-\frac{N-2s}{N}}\Big[a_{N,s}c_{N,s}\phi_a(x)\lambda^{-N+2s} + \mathcal{T}_2(\phi_a(x), \lambda) - a(x)\lambda^{-2s}(\alpha_{N,s} + c_{N,s}d_{N,s}b_{N,s})$$
$$+ \varepsilon\lambda^{-N+2s}Q_V(x)\Big] + o(\lambda^{-2s}) + o(\varepsilon\lambda^{-N+2s}). \tag{2.4}$$

Here, $\mathcal{T}_i(\phi, \lambda)$ are (possibly empty) sums of the form

$$\mathcal{T}_i(\phi, \lambda) := \sum_{k=2}^K \gamma_i(k)\phi^k\lambda^{-k(N-2s)} \tag{2.5}$$

for suitable coefficients $\gamma_i(k) \in \mathbb{R}$, *where* $K = \lfloor \frac{2s}{N-2s} \rfloor$ *is the largest integer less than or equal to* $\frac{2s}{N-2s}$.

Theorem 2.1 is valid irrespective of the criticality of a. The following corollary states two consequences of Theorem 2.1, which concern in particular critical potentials.

Corollary 2.2 (Properties of critical potentials).   *(i) If* $S(a) = S$, *then* $\phi_a(x) \geq 0$ *for all* $x \in \Omega$. *(ii) If* $S(a) = S$ *and* $\phi_a(x) = 0$ *for some* $x \in \Omega$, *then* $a(x) \leq 0$.

Proof. Both statements follow from Theorem 2.1 applied with $\varepsilon = 0$. Indeed, suppose that either $\phi_a(x) < 0$ or $\phi_a(x) = 0 < a(x)$ for some $x \in \Omega$. In both cases, (2.4) gives $S(a) \leq \mathcal{S}_a[\psi_{x,\lambda}] < S$ for $\lambda > 0$ large enough, contradiction. $\qquad\square$

Based on Theorem 2.1, we can now easily derive the following upper bound for $S(a + \varepsilon V)$ provided that $\mathcal{N}_a(V)$ is not empty.

Corollary 2.3. *Suppose that* $\mathcal{N}_a(V) \neq \emptyset$. *Then* $S(a + \varepsilon V) < S$ *for all* $\varepsilon > 0$ *and, as* $\varepsilon \to 0+$,

$$S(a + \varepsilon V) \leq S - \sigma_{N,s} \sup_{x \in \mathcal{N}_a(V)} \frac{|Q_V(x)|^{\frac{2s}{4s-N}}}{|a(x)|^{\frac{N-2s}{4s-N}}} \varepsilon^{\frac{2s}{4s-N}} + o(\varepsilon^{\frac{2s}{4s-N}}), \qquad (2.6)$$

where

$$\sigma_{N,s} = A_{N,s}^{-\frac{N-2s}{N}} (\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s})^{-\frac{N-2s}{4s-N}} \left(\frac{N-2s}{2s} \right)^{\frac{2s}{4s-N}} \frac{4s-N}{N-2s}. \qquad (2.7)$$

Proof. Let us fix $\varepsilon > 0$ and $x \in \mathcal{N}_a(V)$. Then by (2.4)

$$S(a + \varepsilon V) \leq \mathcal{S}_{a+\varepsilon V}[\psi_{x,\lambda}] \qquad\qquad (2.8)$$

$$= S + A_{N,s}^{-\frac{N-2s}{N}} \left(-(a(x) + o(1))\lambda^{-2s}(\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s}) + \varepsilon \lambda^{-N+2s}(Q_V(x) + o(1)) \right). \qquad\qquad (2.9)$$

We first optimize the right side over $\lambda > 0$. Since $A_\varepsilon := (-a(x) + o(1))(\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s})$ and $B_\varepsilon := -Q_V(x) + o(1)$, are strictly positive by our assumptions, we are in the situation of Lemma B.6. Picking $\lambda = \lambda_0(\varepsilon)$ given by (B.5), we have $o(\lambda^{-2s}) = o(\varepsilon^{\frac{2s}{4s-N}})$. Thus, by (B.6), we get, as $\varepsilon \to 0+$,

$$S(a + \varepsilon V) \leq S - \varepsilon^{\frac{2s}{4s-N}} \frac{|Q_V(x)|^{\frac{2s}{4s-N}}}{|a(x)|^{\frac{N-2s}{4s-N}}} A_{N,s}^{-\frac{N-2s}{N}} (\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s})^{-\frac{N-2s}{4s-N}} \left(\frac{N-2s}{2s} \right)^{\frac{2s}{4s-N}} \frac{4s-N}{N-2s}$$

$$+ o(\varepsilon^{\frac{2s}{4s-N}}).$$

Now, optimizing over $x \in \mathcal{N}_a(V)$ completes the proof of (2.6). In particular, $S(a + \varepsilon V) < S$ for small enough $\varepsilon > 0$. Since $S(a + \varepsilon V)$ is a concave function of ε (being the infimum over u of functions $\mathcal{S}_{a+\varepsilon V}[u]$ linear in ε) and $S(a) = S$, this implies that $S(a + \varepsilon V) < S$ for every $\varepsilon > 0$. $\qquad\square$

Proof of Theorem 2.1. **Step 1:** *Expansion of the numerator.* Since $(-\Delta)^s H_a(x, \cdot) = a G_a(x, \cdot)$, the function $\psi_{x,\lambda}$ satisfies

$$(-\Delta)^s \psi_{x,\lambda} = c_{N,s} U_{x,\lambda}^{p-1} - \lambda^{-\frac{N-2s}{2}} a G_a(x, \cdot).$$

Therefore, recalling Lemma A.2,

$$
\begin{aligned}
\|(-\Delta)^{s/2}\psi_{x,\lambda}\|_2^2 &= \int_\Omega \psi_{x,\lambda}(y)(-\Delta)^s \psi_{x,\lambda}(y)\,\mathrm{d}y \\
&= \int_\Omega \left(U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}} H_a(x,\cdot) - f_{x,\lambda}\right)\left(c_{N,s}U_{x,\lambda}^{p-1} - \lambda^{-\frac{N-2s}{2}} aG_a(x,\cdot)\right)\mathrm{d}y \\
&= c_{N,s}\int_\Omega U_{x,\lambda}^p\,\mathrm{d}y - c_{N,s}\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-1}H_a(x,\cdot)\,\mathrm{d}y \\
&\quad - \lambda^{-\frac{N-2s}{2}}\int_\Omega aG_a(x,\cdot)\left(U - \lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)\mathrm{d}y \\
&\quad - \int_\Omega f_{x,\lambda}\left(c_{N,s}U_{x,\lambda}^{p-1} - \lambda^{-\frac{N-2s}{2}}aG_a(x,\cdot)\right)\mathrm{d}y.
\end{aligned}
$$

We now treat the four terms on the right side separately.

A simple computation shows that $\int_{\mathbb{R}^N \setminus B_{d(x)}(x)} U_{x,\lambda}^p\,\mathrm{d}y = \mathcal{O}(\lambda^{-N})$. Thus the first term is given by

$$
c_{N,s}\int_\Omega U_{x,\lambda}^p\,\mathrm{d}y = c_{N,s}\left(\int_{\mathbb{R}^N} U_{0,1}^p\,\mathrm{d}y\right) + \mathcal{O}(\lambda^{-N}) = c_{N,s}A_{N,s} + o(\lambda^{-2s}).
$$

The second term is, by Lemma A.4,

$$
-c_{N,s}\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-1}H_a(x,\cdot)\,\mathrm{d}y = -c_{N,s}a_{N,s}\phi_a(x)\lambda^{-N+2s} + c_{N,s}d_{N,s}b_{N,s}a(x)\lambda^{-2s} + o(\lambda^{-2s}).
$$

The third term will be combined with a term coming from $\int_\Omega(a + \varepsilon V)\psi_{x,\lambda}^2\,\mathrm{d}y$, see below.

The fourth term can be bounded, by Lemma B.1 and recalling $\|f_{x,\lambda}\|_\infty \lesssim \lambda^{-\frac{N+4-2s}{2}}$ from Lemma A.2, by

$$
\left|\int_\Omega f_{x,\lambda}\left(c_{N,s}U_{x,\lambda}^{p-1} - \lambda^{-\frac{N-2s}{2}}aG_a(x,\cdot)\right)\mathrm{d}y\right| \lesssim \|f_{x,\lambda}\|_\infty\left(\|U_{x,\lambda}\|_{p-1}^{p-1} + \lambda^{-\frac{N-2s}{2}}\right) \lesssim \lambda^{-N-2+2s} = o(\lambda^{-2s}).
$$

Now we treat the potential term. We have

$$
\begin{aligned}
\int_\Omega(a + \varepsilon V)\psi_{x,\lambda}^2\,\mathrm{d}y &= \int_\Omega(a + \varepsilon V)\left(U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_a(x,\cdot) - f_{x,\lambda}\right)^2\mathrm{d}y \\
&= \int_\Omega(a + \varepsilon V)\left(U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)^2\mathrm{d}y \\
&\quad - 2\int_\Omega(a + \varepsilon V)f_{x,\lambda}\left(U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)\mathrm{d}y + \int_\Omega(a + \varepsilon V)f_{x,\lambda}^2\,\mathrm{d}y.
\end{aligned}
$$

Similarly to the above, the terms containing $f_{x,\lambda}$ are bounded by

$$
\left|\int_\Omega(a + \varepsilon V)f_{x,\lambda}\left(U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)\mathrm{d}y\right| \lesssim \|f_{x,\lambda}\|_\infty(\|U_{x,\lambda}\|_1 - \lambda^{-\frac{N-2s}{2}}) \lesssim \lambda^{-N-2+2s} = o(\lambda^{-2s})
$$

and

$$
\left|\int_\Omega(a + \varepsilon V)f_{x,\lambda}^2\,\mathrm{d}y\right| \lesssim \|f_{x,\lambda}\|_\infty^2 \lesssim \lambda^{-N-4+2s} = o(\lambda^{-2s}).
$$

Finally, we combine the main term with the third term in the expansion of $\|(-\Delta)^{s/2}\psi_{x,\lambda}\|_2^2$ from above. Recalling that

$$
U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_a(x,\cdot) - \lambda^{-\frac{N-2s}{2}}G_a(x,\cdot) = U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}|x - y|^{-N+2s} = -\lambda^{\frac{N-2s}{2}}h(\lambda(x - y)) \quad (2.10)
$$

with h as in Lemma B.3, we get

$$-\lambda^{-\frac{N-2s}{2}}\int_\Omega aG_a(x,\cdot)\left(U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)\mathrm{d}y + \int_\Omega (a+\varepsilon V)\left(U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)^2\mathrm{d}y$$

$$= -\int_\Omega a\left(U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)\lambda^{\frac{N-2s}{2}}h(\lambda(x-y))\,\mathrm{d}y + \varepsilon\int_\Omega V\left(U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)^2\mathrm{d}y.$$

Since

$$\int_\Omega aH_a(x,\cdot)h(\lambda(x-y))\,\mathrm{d}y \lesssim \lambda^{-N}\|h\|_{L^1(\mathbb{R}^N)} = o(\lambda^{-2s}),$$

by Lemma B.4 we have

$$-\int_\Omega a\left(U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)\lambda^{\frac{N-2s}{2}}h(\lambda(x-y))\,\mathrm{d}y = -\alpha_{N,s}\lambda^{-2s}a(x) + o(\lambda^{-2s}).$$

Moreover, again by (2.10), and using that $h \in L^2(\mathbb{R}^N)$ by Lemma B.3,

$$\varepsilon\int_\Omega V\left(U-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\right)^2$$

$$= \varepsilon\lambda^{-N+2s}\int_\Omega VG_a(x,\cdot)^2\,\mathrm{d}y - 2\varepsilon\int_\Omega VG_a(x,\cdot)h(\lambda(x-y))\,\mathrm{d}y + \varepsilon\lambda^{N-2s}\int_\Omega Vh(\lambda(x-y))^2\,\mathrm{d}y,$$

with

$$\varepsilon\int_\Omega VG_a(x,\cdot)h(\lambda(x-y))\,\mathrm{d}y \lesssim \varepsilon\lambda^{-N/2}\|G_a(x,\cdot)\|_2\|h\|_2 = o(\varepsilon\lambda^{-N+2s})$$

and

$$\varepsilon\lambda^{N-2s}\int_\Omega Vh(\lambda(x-y))^2\,\mathrm{d}y \lesssim \varepsilon\lambda^{-2s}\|h\|_2^2 = o(\varepsilon\lambda^{-N+2s}).$$

This completes the proof of the claimed expansion (2.2).

Step 2: *Expansion of the denominator.* Recall $p = \frac{2N}{N-2s}$. Firstly, writing $\psi_{x,\lambda} = U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_a(x,\cdot) - f_{x,\lambda}$, we have

$$\int_\Omega \psi_{x,\lambda}^p\,\mathrm{d}y = \int_\Omega (U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot))^p\,\mathrm{d}y + \mathcal{O}(\|U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)\|_{p-1}^{p-1}\|f_{x,\lambda}\|_\infty + \|f_{x,\lambda}\|_\infty^p).$$

Using Lemma B.1 and the bound $\|f_{x,\lambda}\|_\infty \lesssim \lambda^{-\frac{N+4-2s}{2}}$, we deduce that the remainder term is $o(\lambda^{-2s})$. To evaluate the main term, from Taylor's formula for $t \mapsto t^p$, we have

$$(a+b)^p = a^p - pa^{p-1}b + \sum_{k=2}^K \binom{p}{k}a^{p-k}b^k + \mathcal{O}(a^{p-K-1}|b|^{K+1} + |b|^p).$$

Here, $\binom{p}{k} := \frac{\Gamma(p+1)}{\Gamma(p-k+1)\Gamma(k+1)}$ is the generalized binomial coefficient and $K = \lfloor\frac{2s}{N-2s}\rfloor$ as in the statement of the theorem. Applying this with $a = U_{x,\lambda}(y)$, $b = -\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot)$, we find

$$\int_\Omega (U_{x,\lambda}-\lambda^{-\frac{N-2s}{2}}H_a(x,\cdot))^p\,\mathrm{d}y = \int_\Omega U_{x,\lambda}^p\,\mathrm{d}y - p\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-1}H_a(x,\cdot)\,\mathrm{d}y$$

$$+ \mathcal{O}\left(\lambda^{-N+2s}\int_\Omega U_{x,\lambda}^{p-2}H_a(x,\cdot)^2\,\mathrm{d}y\right) + \mathcal{O}\left(\lambda^{-N}\int_\Omega H_a(x,\cdot)^p\,\mathrm{d}y\right).$$

By Lemma A.4, the claimed expansion (2.3) follows.

Step 3: *Expansion of the quotient.* For $\alpha = 2/p \in (0,1)$, and fixed $a > 0$, we again use the Taylor expansion

$$(a+b)^{-\alpha} = a^{-\alpha} - \alpha a^{-\alpha-1}b + \sum_{k=2}^K \binom{-\alpha}{k}a^{-\alpha-k}b^k + \mathcal{O}(b^{K+1}).$$

By Step 2, we may apply this with $a = A_{N,s}$ and $b = -p a_{N,s} \phi_a(x) \lambda^{-N+2s} + \mathcal{T}_1(\phi_a(x), \lambda) + p d_{N,s} b_{N,s} a(x) \lambda^{-2s} + o(\lambda^{-2s})$. Since $b = \mathcal{O}(\lambda^{-N+2s})$ and $K + 1 > \frac{2s}{N-2s}$, we have $\mathcal{O}(b^{K+1}) = o(\lambda^{-2s})$ and thus

$$\left(\int_\Omega \psi_{x,\lambda}^p \, \mathrm{d}y \right)^{-2/p} = A_{N,s}^{-\frac{N-2s}{N}} + A_{N,s}^{-\frac{2(N-s)}{N}} \left(2 a_{N,s} \phi_a(x) \lambda^{-N+2s} - 2 d_{N,s} b_{N,s} a(x) \lambda^{-2s} \right) \qquad (2.11)$$

$$+ \mathcal{T}_3(\phi_a(x), \lambda) + o(\lambda^{-2s}), \qquad (2.12)$$

for some term $\mathcal{T}_3(\phi, \lambda)$ as in (2.5). Multiplying this expansion with (2.2), we obtain

$$\mathcal{S}_{a+\varepsilon V}[\psi_{x,\lambda}] = c_{n,s} A_{N,s}^{\frac{2s}{N}} + A_{N,s}^{-\frac{N-2s}{N}} \left[a_{N,s} c_{N,s} \phi_a(x) \lambda^{-N+2s} + \mathcal{T}_2(\phi_a(x), \lambda) \right.$$

$$\left. - a(x) \lambda^{-2s} (\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s}) + \varepsilon \lambda^{-N+2s} Q_V(x) \right]$$

$$+ o(\lambda^{-2s}) + o(\varepsilon \lambda^{-N+2s}).$$

By integrating the equation $(-\Delta)^s U_{0,1} = c_{N,s} U_{0,1}^{p-1}$ and using the fact that $U_{0,1}$ minimizes the Sobolev quotient on $\mathbb{R}^N$ (or by a computation on the numerical values of the constants given in Lemma B.5), we have $c_{N,s} A_{N,s}^{\frac{2s}{N}} = S$. Hence, this is the expansion claimed in (2.4). $\qquad \square$

3. Proof of the lower bound I: a first expansion

3.1. **Non-existence of a minimizer for $S(a)$.** In this section, we prove that for a critical potential a, the infimum $S(a)$ is not attained. As we will see in Section 3.2, this implies the important basic fact that the minimizers for $S(a + \varepsilon V)$ must blow up as $\varepsilon \to 0$.

The following is the main result of this section.

Proposition 3.1 (Non-existence of a minimizer for $S(a)$). *Suppose that $a \in C(\overline{\Omega})$ is a critical potential. Then*

$$S(a) = \inf_{u \in \widetilde{H}^s(\Omega), u \not\equiv 0} \frac{\int_\Omega |(-\Delta)^{s/2} u|^2 \, \mathrm{d}y + \int_\Omega a u^2 \, \mathrm{d}y}{\|u\|_{\frac{2N}{N-2s}}^2}$$

is not achieved.

For $s = 1$, Proposition 3.1 was proved by Druet [20] and we follow his strategy. The feature that makes the generalization of [20] to $s \in (0,1)$ not completely straightforward is its use of the product rule for ordinary derivatives. Instead, we shall use the identity

$$(-\Delta)^s (uv) = u(-\Delta)^s v + v(-\Delta)^s u - I_s(u,v), \qquad (3.1)$$

where

$$I_s(u,v)(x) := C_{N,s} P.V. \int_{\mathbb{R}^N} \frac{(u(x) - u(y))(v(x) - v(y))}{|x - y|^{N+2s}} \, \mathrm{d}y,$$

with $C_{N,s}$ as in (1.6). While the relation (3.1) can be verified by a simple computation (see e.g. [25, Lemma 20.2]), it leads to more complicated terms than those arising in Druet's proof. To be more precise, the term $\int_\Omega u^2 |\nabla \varphi|^2$ from [20] is replaced by the term $\mathcal{I}(\varphi)$ defined in (3.6), which is more involved to evaluate for the right choice of φ.

Proof of Proposition 3.1. For the sake of finding a contradiction, we suppose that there exists u which achieves $S(a)$, normalized so that

$$\int_\Omega u^{\frac{2N}{N-2s}} \, \mathrm{d}y = 1. \qquad (3.2)$$

Then u satisfies the equation

$$(-\Delta)^s u + a u = S u^{\frac{N-2s}{N-2s}}, \qquad (3.3)$$

with Lagrange multiplier $S = S_{N,s}$ equal to the Sobolev constant. (Indeed, this value is determined by integrating the equation against u and using (3.2).)

Since $S(a) = S$, we have, for every $\varphi \in C^\infty(\mathbb{R}^N)$ and $\varepsilon > 0$, and abbreviating $p = \frac{2N}{N-2s}$,

$$S \left(\int_\Omega (u(1 + \varepsilon\varphi))^p \, \mathrm{d}y \right)^{\frac{2}{p}} \leq \int_{\mathbb{R}^N} |(-\Delta)^{s/2} (u(1 + \varepsilon\varphi))|^2 \, \mathrm{d}y + \int_\Omega au^2(1 + \varepsilon\varphi)^2 \, \mathrm{d}y. \qquad (3.4)$$

We shall expand both sides of (3.4) in powers of ε. For the left side, a simple Taylor expansion together with (3.2) gives

$$\left(\int_\Omega (u(1 + \varepsilon\varphi))^p \, \mathrm{d}y \right)^{\frac{2}{p}} = 1 + 2\varepsilon \int_\Omega u^p\varphi \, \mathrm{d}y + \varepsilon^2 \left((p-1) \int_\Omega u^p\varphi^2 \, \mathrm{d}y - (p-2) \left(\int_\Omega u^p\varphi \, \mathrm{d}y \right)^2 \right) + o(\varepsilon^2).$$
$$(3.5)$$

Expanding the right side is harder and we need to invoke the fractional product rule (3.1). Firstly, integrating by parts we have

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2} (u(1 + \varepsilon\varphi))|^2 \, \mathrm{d}y = \int_{\mathbb{R}^N} u(1 + \varepsilon\varphi)(-\Delta)^s (u(1 + \varepsilon\varphi)) \, \mathrm{d}y.$$

By (3.1), we can write

$$(-\Delta)^s (u(1 + \varepsilon\varphi)) = (1 + \varepsilon\varphi)(-\Delta)^s u + \varepsilon u(-\Delta)^s\varphi - \varepsilon I_s(u, \varphi).$$

Hence

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2} (u(1 + \varepsilon\varphi))|^2 \, \mathrm{d}y = \int_\Omega u(-\Delta)^s u(1 + \varepsilon\varphi)^2 + \varepsilon^2 \mathcal{I}(\varphi),$$

where we write

$$\mathcal{I}(\varphi) := \varepsilon^{-1} \int_\Omega u(1 + \varepsilon\varphi) \left(u(-\Delta)^s\varphi - I_s(u, \varphi) \right) \mathrm{d}y. \qquad (3.6)$$

Writing out $(-\Delta)^s\varphi$ as the singular integral given by (1.5), we obtain (we drop the principal value for simplicity)

$$\mathcal{I}(\varphi) = \varepsilon^{-1} C_{N,s} \iint_{\mathbb{R}^N \times \mathbb{R}^N} u(x)u(y)(1 + \varepsilon\varphi(x)) \frac{\varphi(x) - \varphi(y)}{|x - y|^{N+2s}} \, \mathrm{d}x \, \mathrm{d}y \qquad (3.7)$$

$$= \frac{C_{N,s}}{2} \iint_{\mathbb{R}^N \times \mathbb{R}^N} u(x)u(y) \frac{|\varphi(x) - \varphi(y)|^2}{|x - y|^{N+2s}} \, \mathrm{d}x \, \mathrm{d}y.$$

The last equality follows by symmetrizing in the x and y variables.

Thus we can write the right side of (3.4) as

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2} (u(1 + \varepsilon\varphi))|^2 \, \mathrm{d}y + \int_\Omega au^2(1 + \varepsilon\varphi)^2 \, \mathrm{d}y$$

$$= \int_\Omega u((-\Delta)^s u + au)(1 + \varepsilon\varphi)^2 \, \mathrm{d}y + \varepsilon^2 \mathcal{I}(\varphi)$$

$$= S \int_\Omega u^p(1 + \varepsilon\varphi)^2 \, \mathrm{d}y + \varepsilon^2 \mathcal{I}(\varphi),$$

where we used equation (3.3). After expanding the square $(1 + \varepsilon\varphi)^2$, the terms of orders 1 and ε on both sides of (3.4) cancel precisely. For the coefficients of ε^2, we thus recover the inequality

$$\int_\Omega u^p\varphi^2 \, \mathrm{d}y \leq \frac{1}{S(p-2)} \mathcal{I}(\varphi) + \left(\int_\Omega u^p\varphi \, \mathrm{d}y \right)^2. \qquad (3.8)$$

We now make a suitable choice of φ, which turns (3.8) into the desired contradiction. As in [20], we choose

$$\varphi_i(y) := (\mathcal{S}(y))_i, \quad i = 1, ..., N + 1,$$

where $\mathcal{S} : \mathbb{R}^N \to \mathbb{S}^N$ is the (inverse) stereographic projection, i.e. [31, Sec. 4.4]

$$\varphi_i = \frac{2y_i}{1 + |y|^2} \quad \text{for } i = 1, ..., N, \qquad \varphi_{N+1} = \frac{1 - |y|^2}{1 + |y|^2}. \tag{3.9}$$

Moreover we may assume (up to scaling and translating Ω if necessary) that the balancing condition

$$\int_\Omega u^p \varphi_i \, \mathrm{d}y = 0, \qquad i = 1, ..., N + 1 \tag{3.10}$$

is satisfied. Since [20] is rather brief on this point, we include some details in Lemma 3.2 below for the convenience of the reader.

By definition, we have $\sum_{i=1}^{N+1} \varphi_i^2 = 1$. Testing (3.8) with φ_i and summing over i thus yields, by (3.10),

$$1 = \int_\Omega u^p \, \mathrm{d}y \leq \frac{1}{S(p-2)} \sum_{i=1}^{N+1} \mathcal{I}(\varphi_i). \tag{3.11}$$

To obtain a contradiction and finish the proof, we now show that $\sum_{i=1}^{N+1} \mathcal{I}(\varphi_i) < S(p-2)$. By definition of φ_i, we have

$$\sum_{i=1}^{N+1} \mathcal{I}(\varphi_i) = \frac{C_{N,s}}{2} \iint_{\mathbb{R}^N \times \mathbb{R}^N} u(x)u(y) \frac{|\mathcal{S}(x) - \mathcal{S}(y)|^2}{|x-y|^{N+2s}} \, \mathrm{d}x \, \mathrm{d}y. \tag{3.12}$$

To evaluate this integral further, we pass to $\mathbb{S}^N$. Set $J_{\mathcal{S}^{-1}}(\eta) := \det D\mathcal{S}^{-1}(\eta)$ and define

$$U(\eta) := u(\mathcal{S}^{-1}(\eta))J_{\mathcal{S}^{-1}}(\eta)^{\frac{1}{p}},$$

so that $\int_{\mathbb{S}^N} U^p \, \mathrm{d}\eta = 1$. Since the distance transforms as

$$|\mathcal{S}^{-1}(\eta) - \mathcal{S}^{-1}(\xi)| = J_{\mathcal{S}^{-1}}(\eta)^{\frac{1}{2N}} |\eta - \xi| J_{\mathcal{S}^{-1}}(\xi)^{\frac{1}{2N}},$$

changing variables in (3.12) gives

$$\sum_{i=1}^{N+1} \mathcal{I}(\varphi_i) = \frac{C_{N,s}}{2} \iint_{\mathbb{S}^N \times \mathbb{S}^N} \frac{U(\eta)U(\xi)}{|\eta - \xi|^{N+2s-2}} \, \mathrm{d}\eta \, \mathrm{d}\xi \tag{3.13}$$

By applying first Cauchy–Schwarz and then Hölder's inequality, we estimate

$$\sum_{i=1}^{N+1} I(\varphi_i) \leq \frac{C_{N,s}}{2} \iint_{\mathbb{S}^N \times \mathbb{S}^N} \frac{U(\eta)^2}{|\eta - \xi|^{N+2s-2}} \, \mathrm{d}\eta \, \mathrm{d}\xi$$

$$= \frac{C_{N,s}}{2} \delta_{N,s} \int_{\mathbb{S}^N} U(\eta)^2 \, \mathrm{d}\eta < \frac{C_{N,s}}{2} \delta_{N,s} |\mathbb{S}^N|^{\frac{2s}{N}}, \tag{3.14}$$

where the last inequality is strict. Indeed, U vanishes near the south pole of $\mathbb{S}^N$, hence there cannot be equality in Hölder's inequality applied with the functions U^2 and 1. Moreover, in the above we abbreviated

$$\delta_{N,s} := \int_{\mathbb{S}^N} \frac{1}{|\eta - \xi|^{N+2s-2}} \, \mathrm{d}\xi$$

(note that this number is independent of $\eta \in \mathbb{S}^N$). By transforming back to $\mathbb{R}^n$ and evaluating a Beta function integral, the explicit value of $\delta_{N,s}$ can be computed explicitly to be

$$\delta_{N,s} = 2^{2-2s} \pi^{N/2} \frac{\Gamma(1-s)}{\Gamma(\frac{N}{2} + 1 - s)}.$$

Inserting this into estimate (3.14), as well as the explicit values of $C_{N,s}$ given in (1.6) and of $S_{N,s}$ given in (1.9), a direct computation then gives

$$\frac{\sum_{i=1}^{N+1} \mathcal{I}(\varphi_i)}{S_{N,s}} < \frac{1}{S_{N,s}} \frac{C_{N,s}}{2} \delta_{N,s} |\mathbb{S}^N|^{\frac{2s}{N}} = \frac{s2^{2-2s}}{N-2s} \left(2 \frac{\Gamma(N)\Gamma(\frac{1}{2})}{\Gamma(\frac{N+1}{2})\Gamma(\frac{N}{2})} \right)^{\frac{2s}{N}}.$$

It can be easily shown by induction over N that

$$2\frac{\Gamma(N)\Gamma(\frac{1}{2})}{\Gamma(\frac{N+1}{2})\Gamma(\frac{N}{2})} = 2^N$$

for every $N \in \mathbb{N}$, and hence

$$\sum_{i=1}^{N+1} \mathcal{I}(\varphi_i) < S\frac{4s}{N-2s} = S(p-2).$$

This is the desired contradiction to (3.11). $\square$

Here is the lemma that we referred to in the previous proof. It expands an argument sketched in [20, Step 1]. To emphasize its generality, instead of u^p we state it for a general nonnegative function h with $\int_\Omega h = 1$.

Lemma 3.2. *Let $\Omega \subset \mathbb{R}^N$ be an open bounded set and $0 \le h \in L^1(\Omega)$ with $\|h\|_1 = 1$. Then there is $(y,t) \in \mathbb{R}^N \times (0,\infty)$ such that*

$$F(y,t) := \int_\Omega h(x)\frac{2t(x-y)}{1+t^2|x-y|^2}\,\mathrm{d}x = 0,$$

$$G(y,t) := \int_\Omega h(x)\frac{1-t^2|x-y|^2}{1+t^2|x-y|^2}\,\mathrm{d}x = 0.$$

Proof. Define $H : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^{N+1}$ by

$$H(y,s) := \left(F\left(y,\frac{s+\sqrt{s^2+4}}{2}\right)+y, G\left(y,\frac{s+\sqrt{s^2+4}}{2}\right)+s\right).$$

We claim that

$$|H(y,s)| \le |y|^2 + s^2 \qquad (3.15)$$

whenever $|y|^2+s^2$ is large enough. Thus for large enough radii $R > 0$, the map H sends $B(0,R) \subset \mathbb{R}^{N+1}$ into itself. By the Brouwer fixed point theorem, H has a fixed point (y,s). Then the pair $(y,\frac{s+\sqrt{s^2+4}}{2})$ satisfies the property stated in the lemma.

To prove (3.15), it is more natural to set $t := \frac{s+\sqrt{s^2+4}}{2} > 0$, so that $s = t - t^{-1}$. By writing out $|H(y,s)|^2$, (3.15) is equivalent to

$$2y \cdot F(y,t) + 2(t-t^{-1})G(y,t) + |F(y,t)|^2 + |G(y,t)|^2 \le 0 \qquad (3.16)$$

whenever $|y|^2 + (t-t^{-1})^2$ is large enough.

First, it is easy to see that $y \cdot F(y,t)$, $F(y,t)$ and $G(y,t)$ are bounded in absolute value uniformly in $(y,t) \in \mathbb{R}^N \times (0,\infty)$. Moreover, there is $C > 0$ such that

$$G(y,t) = 1 - 2\int_\Omega h(x)\frac{t^2|x-y|^2}{1+t^2|x-y|^2}\,\mathrm{d}x \begin{cases} \ge \frac{1}{2} & \text{if } 0 < t \le 1/C, \\ \le -\frac{1}{2} & \text{if } t \ge C. \end{cases}$$

Therefore $(t-t^{-1})G(y,t) \to -\infty$ as $t \to 0$ or $t \to \infty$. Thus (3.16) holds whenever $(t-t^{-1})^2$ is large enough.

Thus we assume in the following that $t \in [1/C, C]$ and prove that (3.16) holds if $|y|$ is large enough. For convenience, fix some sequence (y,t) with $|y| \to \infty$ and $t \to t_0 \in (0,\infty)$. Then $|F(y,t)| \to 0$ and $G(y,t) \to -1$. Moreover, since Ω is bounded, $\frac{|x-y|}{|y|} \to 1$ uniformly in $x \in \Omega$ and hence

$$2y \cdot F(y,t) = -\int_\Omega h(x)\frac{4t|y|^2}{1+t^2|x-y|^2} + \mathcal{O}\left(\int_\Omega h(x)\frac{t|y|}{1+t^2|x-y|^2}\,\mathrm{d}x\right) \to -\frac{4}{t_0}.$$

Altogether, the quantity on the left side of (3.16) thus tends to $-2t_0 - 2t_0^{-1} + 1 \leq -3 < 0$, which concludes the proof of (3.16). $\qquad\square$

3.2. **Profile decomposition.** The following proposition gives an asymptotic decomposition of a general sequence of normalized (almost) minimizers of $S(a + \varepsilon V)$.

Proposition 3.3 (Profile decomposition). *Let $a \in C(\overline{\Omega})$ be critical and let $V \in C(\overline{\Omega})$ be such that $\mathcal{N}_a(V) \neq \emptyset$. Suppose that $(u_\varepsilon) \subset \widetilde{H}^s(\Omega)$ is a sequence such that*

$$\lim_{\varepsilon \to 0} \frac{\mathcal{S}_{a+\varepsilon V}[u_\varepsilon] - S(a + \varepsilon V)}{S - S(a + \varepsilon V)} = 0 \quad and \quad \int_\Omega u_\varepsilon^p \, \mathrm{d}y = \int_{\mathbb{R}^N} U_{0,1}^p \, \mathrm{d}y. \tag{3.17}$$

Then there are sequences $(x_\varepsilon) \subset \Omega$, $(\lambda_\varepsilon) \subset (0, \infty)$, $(w_\varepsilon) \subset T_{x_\varepsilon, \lambda_\varepsilon}^\perp$, and $(\alpha_\varepsilon) \subset \mathbb{R}$ such that, up to extraction of a subsequence,

$$u_\varepsilon = \alpha_\varepsilon \left(P U_{x_\varepsilon, \lambda_\varepsilon} + w_\varepsilon \right). \tag{3.18}$$

Moreover, as $\varepsilon \to 0$, we have

$$\|(-\Delta)^{s/2} w_\varepsilon\|_2 \to 0,$$
$$d(x_\varepsilon)\lambda_\varepsilon \to \infty,$$
$$x_\varepsilon \to x_0,$$
$$\alpha_\varepsilon \to \pm 1.$$

In all of the following, we shall always work with a sequence u_ε that satisfies the assumptions of Proposition 3.3. For readability, we shall often drop the index ε from α_ε, x_ε, λ_ε and w_ε, and write $d := d_\varepsilon := d(x_\varepsilon)$. Moreover, we make the convention that we always assume the strict inequality

$$S(a + \varepsilon V) < S. \tag{3.19}$$

In Theorems 1.4 and 1.6 we assume $\mathcal{N}_a(V) \neq \emptyset$, so assumption (3.19) is certainly justified in view of Corollary 2.3. For Theorem 1.5, where we assume $\mathcal{N}_a(V) = \emptyset$, we discuss the role of assumption (3.19) in the proof of that theorem in Section 6.

Proof. **Step 1.** We derive a preliminary decomposition in terms of the Sobolev optimizers $U_{z,\lambda}$ and without orthogonality condition on the remainder, see (3.20) below.

The assumptions imply that the sequence (u_ε) is bounded in $\widetilde{H}^s(\Omega)$, hence up to a subsequence we may assume $u_\varepsilon \rightharpoonup u_0$ for some $u_0 \in \widetilde{H}^s(\Omega)$. By the argument given in [24, Proof of Proposition 3.1, Step 1], the fact that $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon] \to S(a) = S$ implies that either u_0 is a minimizer for $S(a)$, unless $u_0 \equiv 0$. Since such a minimizer does not exist by Proposition 3.1, we conclude that in fact $u_\varepsilon \rightharpoonup 0$ in $\widetilde{H}^s(\Omega)$.

By Rellich's theorem, $u_\varepsilon \to 0$ strongly in $L^2(\Omega)$, in particular $\int_\Omega (a + \varepsilon V) u_\varepsilon^2 = o(1)$. The assumption (1.17) therefore implies that (u_ε) is a minimizing sequence for the Sobolev quotient $\int_{\mathbb{R}^N} |(-\Delta)^{s/2} u|^2 \, \mathrm{d}y / \|u\|_{\frac{2N}{N-2s}}^2$. Therefore the assumptions of [32, Theorem 1.3] are satisfied, and we may conclude by that theorem that there are sequences $(z_\varepsilon) \subset \mathbb{R}^N$, $(\mu_\varepsilon) \subset (0, \infty)$, (σ_ε) such that

$$\mu_\varepsilon^{-\frac{N-2s}{2}} u_\varepsilon(z_\varepsilon + \mu_\varepsilon^{-1} \cdot) \to \beta U_{0,1}$$

in $\dot{H}^s(\mathbb{R}^N)$, for some $\beta \in \mathbb{R}$. By the normalization condition from (1.17), $\beta \in \{\pm 1\}$. Now, a change of variables $y = z_\varepsilon + \mu_\varepsilon^{-1} x$ implies

$$u_\varepsilon(y) = U_{z_\varepsilon, \mu_\varepsilon}(y) + \sigma_\varepsilon, \tag{3.20}$$

where still $\sigma_\varepsilon \to 0$ in $\dot{H}^s(\mathbb{R}^N)$, since the $\dot{H}^s(\mathbb{R}^N)$-norm is invariant under this change of variable.

Moreover, since Ω is smooth, the fact that

$$\int_{\mu_\varepsilon\Omega+z_\varepsilon} U_{0,1}^p \, \mathrm{d}y = \int_\Omega U_{z_\varepsilon,\mu_\varepsilon}^p \, \mathrm{d}y = \int_{\mathbb{R}^N} U_{0,1}^p \, \mathrm{d}y + o(1)$$

implies $\mu_\varepsilon \mathrm{dist}(z_\varepsilon, \mathbb{R}^N \setminus \Omega) \to \infty$.

Step 2. We make the necessary modifications to derive (3.18) from (3.20). The crucial argument is furnished by [2, Proposition 4.3], which generalizes the corresponding statement by Bahri and Coron [3, Proposition 7] to fractional $s \in (0,1)$. It states the following. Suppose that $u \in \widetilde{H}^s(\Omega)$ with $\|u\|_{\widetilde{H}^s(\Omega)} = A_{N,s}$ satisfies

$$\inf\left\{ \|(-\Delta)^{s/2}(u - PU_{x,\lambda})\|_2 \, : \, x \in \Omega, \, \lambda d(x) > \eta^{-1} \right\} < \eta. \tag{3.21}$$

for some $\eta > 0$. Then if η is small enough, the minimization problem

$$\inf\left\{ \|(-\Delta)^{s/2}(u - \alpha PU_{x,\lambda})\|_2 \, : \, x \in \Omega, \, \lambda d(x) > (4\eta)^{-1}, \, \alpha \in (1/2, 2) \right\} \tag{3.22}$$

has a unique solution.

By the decomposition from Step 1 and Lemma A.2, we have

$$\|(-\Delta)^{s/2}(u_\varepsilon - PU_{z_\varepsilon,\mu_\varepsilon})\|_2 \leq \|(-\Delta)^{s/2}(U_{z_\varepsilon,\mu_\varepsilon} - PU_{z_\varepsilon,\mu_\varepsilon})\|_2 + \|(-\Delta)^{s/2}\sigma_\varepsilon\|_2 \to 0$$

as $\varepsilon \to 0$, so that (3.21) is satisfied by u_ε for all ε small enough, with a constant η_ε tending to zero. We thus obtain the desired decomposition

$$u_\varepsilon = \alpha_\varepsilon(PU_{x_\varepsilon,\lambda_\varepsilon} + w_\varepsilon)$$

by taking $(x_\varepsilon, \lambda_\varepsilon, \alpha_\varepsilon)$ to be the solution to (3.22) and $w_\varepsilon := \alpha_\varepsilon^{-1}u_\varepsilon - PU_{x_\varepsilon,\lambda_\varepsilon}$. To verify the claimed asymptotic behavior of the parameters, note that since $\eta_\varepsilon \to 0$, by definition of the minimization problem (3.22), we have $\|(-\Delta)^{s/2}w_\varepsilon\|_2 < \eta_\varepsilon \to 0$ and $\lambda_\varepsilon d(x_\varepsilon) > (4\eta_\varepsilon)^{-1} \to \infty$. Since Ω is bounded, the convergence $x_\varepsilon \to x_0 \in \overline{\Omega}$ is ensured by passing to a suitable subsequence. Finally, using (1.17) we have

$$\int_{\mathbb{R}^N} U_{0,1}^p \, \mathrm{d}y = \int_\Omega u_\varepsilon^p \, \mathrm{d}y = |\alpha_\varepsilon|^p \int_\Omega PU_{x_\varepsilon,\lambda_\varepsilon} \, \mathrm{d}y + o(1) = |\alpha_\varepsilon|^p \int_{\mathbb{R}^N} U_{0,1}^p \, \mathrm{d}y + o(1),$$

which implies $\alpha_\varepsilon = \pm 1 + o(1)$. □

3.3. **Coercivity.** In the following sections, our goal is to improve the bounds from Proposition 3.3 step by step.

The following inequality, and its improvement in Proposition 3.5 below, will be central. For $s = 1$, these inequalities are due to Rey [34, Eq. (D.1)] and Esposito [21, Lemma 2.1], respectively, whose proofs inspired those given below.

Proposition 3.4 (Coercivity inequality). *For all $x \in \mathbb{R}^n$ and $\lambda > 0$, we have*

$$\|(-\Delta)^{s/2}v\|_2^2 - c_{N,s}(p-1)\int_\Omega U_{x,\lambda}^{p-2}v^2 \, \mathrm{d}y \geq \frac{4s}{N+2s+2}\|(-\Delta)^{s/2}v\|_2^2, \tag{3.23}$$

for all $v \in T_{x,\lambda}^\perp$.

As a corollary, we can include the lower order term $\int_\Omega av^2$, at least when $d(x)\lambda$ is large enough and at the price of having a non-explicit constant on the right side. This is the form of the inequality which we shall use below to refine our error bounds in Sections 3.4 and 5.2.

Proposition 3.5 (Coercivity inequality with potential a). *Let $(x_n) \subset \Omega$ and $(\lambda_n) \subset (0,\infty)$ be sequences such that $\mathrm{dist}(x_n, \partial\Omega)\lambda_n \to \infty$. Then there is $\rho > 0$ such that for all n large enough,*

$$\|(-\Delta)^{s/2}v\|_2^2 + \int_\Omega av^2 \, \mathrm{d}y - c_{N,s}(p-1)\int_\Omega U_{x_n,\lambda_n}^{p-2}v^2 \, \mathrm{d}y \geq \rho\|(-\Delta)^{s/2}v\|_2^2, \qquad \textit{for all } v \in T_{x_n,\lambda_n}^\perp. \tag{3.24}$$

Proof. Abbreviate $U_n := U_{x_n, \lambda_n}$ and $T_n := T_{x_n, \lambda_n}$. We follow the proof of [21] and define

$$C_n := \inf \left\{ 1 + \int_\Omega av^2 \, dy - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v^2 \, dy \, : \, v \in T_n^\perp, \, \|(-\Delta)^{s/2} v\| = 1 \right\}.$$

Then C_n is bounded from below, uniformly in n. We first claim that C_n is achieved whenever $C_n < 1$. Indeed, fix n and let v_k be a minimizing sequence. Up to a subsequence, $v_k \rightharpoonup v_\infty$ in $\widetilde{H}^s(\Omega)$ and consequently $\|(-\Delta)^{s/2} v_\infty\| \leq 1$ and $\int_\Omega av_k^2 - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v_k^2 \, dy \to \int_\Omega av_\infty^2 - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v_\infty^2 \, dy$, by compact embedding $\widetilde{H}^s(\Omega) \hookrightarrow L^2(\Omega)$. Thus

$$(1 - C_n) \|(-\Delta)^{s/2} v_\infty\|^2 + \int_\Omega av_\infty^2 \, dy - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v_\infty^2 \, dy$$

$$\leq (1 - C_n) + \int_\Omega av_\infty^2 \, dy - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v_\infty^2 \, dy = 0.$$

On the other hand, the left hand side of the above inequality must itself be non-negative, for otherwise $\tilde{v} := v_\infty / \|(-\Delta)^{s/2} v_\infty\|$ (notice that $C_n < 1$ enforces $v_\infty \not\equiv 0$) yields a contradiction to the definition of C_n as an infimum. Thus the above inequality must be in fact an equality, whence $\|(-\Delta)^{s/2} v_\infty\| = 1$. We have thus proved that C_n is achieved if $C_n < 1$.

Now, assume for contradiction, up to passing to a subsequence, that $\lim_{n\to\infty} C_n =: L \leq 0$. By the first part of the proof, let v_n be a minimizer satisfying

$$(1 - C_n) \int_{\mathbb{R}^N} (-\Delta)^{s/2} v_n (-\Delta)^{s/2} w \, dy + \int_\Omega av_n w \, dy - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v_n w \, dy = 0 \qquad (3.25)$$

for all $w \in T_n^\perp$. Up to passing to a subsequence, we may assume $v_n \rightharpoonup v \in \widetilde{H}^s(\Omega)$. We claim that

$$(1 - L)(-\Delta)^s v + av = 0 \qquad \text{in } (\widetilde{H}^s(\Omega))'. \tag{3.26}$$

Assuming (3.26) for the moment, we obtain a contradiction as follows. Testing (3.26) against $v \in \widetilde{H}^s(\Omega)$ gives

$$\|(-\Delta)^{s/2} v\|^2 + \int_\Omega av^2 \, dy = L \|(-\Delta)^{s/2} v\|^2 \leq 0.$$

On the other hand, by coercivity of $(-\Delta)^s + a$, the left hand side must be nonnegative and hence $v \equiv 0$. By compact embedding, we deduce $v_n \to 0$ strongly in $L^2(\Omega)$ and thus

$$C_n = 1 - c_{N,s}(p-1) \int_\Omega U_n^2 v_n^2 \, dy + o(1) \geq \frac{4s}{N + 2s + 2} + o(1).$$

This is the desired contradiction to $\lim_{n\to\infty} C_n \leq 0$.

At last, we prove (3.26). Let $\varphi \in \widetilde{H}^s(\Omega)$ be given and write $\varphi = u_n + w_n$, with $u_n \in T_n$ and $w_n \in T_n^\perp$. By (3.25) and using $\int_{\mathbb{R}^N} (-\Delta)^{s/2} v_n (-\Delta)^{s/2} u_n = 0$,

$$(1 - C_n) \int_{\mathbb{R}^N} (-\Delta)^{s/2} v_n (-\Delta)^{s/2} \varphi \, dy + \int_\Omega av_n \varphi \, dy - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v_n \varphi \, dy \qquad (3.27)$$

$$= \int_\Omega av_n u_n \, dy - c_{N,s}(p-1) \int_\Omega U_n^{p-2} v_n u_n \, dy = \mathcal{O}(\|(-\Delta)^{s/2} u_n\|). \tag{3.28}$$

On the one hand, we have

$$\left| \int_\Omega U_n^{p-2} v_n \varphi \, dy \right| \leq \|U_n^{p-2} \varphi\|_{\frac{p}{p-1}} \to 0$$

because $\varphi^{\frac{p}{p-1}} \in L^{p-1} = (L^{\frac{p-1}{p-2}}(\Omega))'$ and $U_n^{\frac{(p-2)p}{p-1}} \rightharpoonup 0$ weakly in $L^{\frac{p-1}{p-2}}(\Omega)$. Thus, by weak convergence, the expression in (3.27) tends to

$$(1 - L) \int_{\mathbb{R}^N} (-\Delta)^{s/2} v (-\Delta)^{s/2} \varphi \, dy + \int_\Omega av\varphi \, dy$$

as desired. In view of (3.28), the proof of (3.26) is thus complete if we can show $\|(-\Delta)^{s/2}u_n\| \to 0$. This is again a consequence of weak convergence to zero of the U_n. Indeed, by Lemmas B.1 and B.2 we have

$$\left|\int_{\mathbb{R}^N} (-\Delta)^{s/2} \frac{PU_n}{\|(-\Delta)^{s/2}PU_n\|}(-\Delta)^{s/2}\varphi\,\mathrm{d}y\right| \lesssim \int_{\mathbb{R}^N} U_n^{p-1}|\varphi|\,\mathrm{d}y = o(1),$$

$$\left|\int_{\mathbb{R}^N} (-\Delta)^{s/2} \frac{\partial_\lambda PU_n}{\|(-\Delta)^{s/2}\partial_\lambda PU_n\|}(-\Delta)^{s/2}\varphi\,\mathrm{d}y\right| \lesssim \lambda \int_{\mathbb{R}^N} U_n^{p-2}\partial_\lambda U_n\varphi\,\mathrm{d}y \lesssim \|U_n^{p-2}\varphi\|_{\frac{p}{p-1}} = o(1),$$

and similarly

$$\int_{\mathbb{R}^N} (-\Delta)^{s/2}(\lambda^{-N+2s}\partial_{x_i}PU_n)(-\Delta)^{s/2}\varphi\,\mathrm{d}y = o(1).$$

Here we used again that $U_n^{p-1} \rightharpoonup 0$ in $L^{\frac{p}{p-1}}$ and $U_n^{\frac{(p-2)p}{p-1}} \rightharpoonup 0$ in $L^{\frac{p-1}{p-2}}$ weakly. $\int_\Omega U^{\frac{(p-2)p}{p-1}}\varphi^{\frac{p}{p-1}}\,\mathrm{d}y = o(1)$ by weak convergence.

From the convergence to zero of these scalar products, one can conclude $u_n \to 0$ by using the fact that the PU_n, $\partial_\lambda PU_n$, $\partial_{x_i}PU_n$ are 'asymptotically orthogonal' by the bounds of Lemma B.2. For a detailed argument, we refer to Lemma 5.2 below, see also [24, Lemma 6.1]. $\qquad\square$

Let us now prepare the proof of Proposition 3.4. We recall that $\mathcal{S}:\mathbb{R}^N \to \mathbb{S}^N\setminus\{S\}$ (where $S = -e_{N+1}$ is the southpole) denotes the inverse stereographic projection defined in (3.9), with Jacobian $J_\mathcal{S}(x) := \det D\mathcal{S}(x) = \left(\frac{2}{1+|x|^2}\right)^N$.

Given a function v on $\mathbb{R}^N$, we may define a function u on $\mathbb{S}^N$ by setting

$$u(\omega) := v_{\mathcal{S}^{-1}}(\omega) := v(\mathcal{S}^{-1}(\omega))J_{\mathcal{S}^{-1}}(\omega)^{\frac{N-2s}{2N}}, \qquad \omega \in \mathbb{S}^N\setminus\{S\}.$$

The inverse of this map is of course given by

$$v(y) := u_\mathcal{S}(y) := u(\mathcal{S}(y))J_\mathcal{S}(y)^{\frac{N-2s}{2N}}, \qquad y \in \mathbb{R}^N.$$

The exponent in the determinantal factor is chosen such that $\|v\|_{L^p(\mathbb{R}^N)} = \|u\|_{L^p(\mathbb{S}^N)}$.

For a basis $(Y_{l,m})$ of $L^2(\mathbb{S}^N)$ consisting of L^2-normalized spherical harmonics, write $u \in L^2(\mathbb{S}^N)$ as $u = \sum_{l,m} u_{l,m}Y_{l,m}$ with coefficients $u_{l,m} \in \mathbb{R}$. With

$$\lambda_l = \frac{\Gamma(l + \frac{N}{2} + s)}{\Gamma(l + \frac{N}{2} - s)}. \tag{3.29}$$

the Paneitz operator $\mathcal{P}_{2s}$ is defined by

$$\mathcal{P}_{2s}u := \sum_{l,m} \lambda_l u_{l,m}Y_{l,m}$$

for every $u \in L^2(\mathbb{S}^N)$ such that $\sum_{l,m}\lambda_l u_{l,m}^2 < \infty$.

It is well-known (see [5]) that, for every $v \in C_0^\infty(\mathbb{R}^N)$, we have,

$$(-\Delta)^s v(x) = J_\mathcal{S}(x)^{\frac{N+2s}{2N}}\mathcal{P}_{2s}u(\mathcal{S}(x)), \tag{3.30}$$

where $u = v_\mathcal{S}$. Thus we have

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2}v|^2\,\mathrm{d}y = \int_{\mathbb{R}^N} v(-\Delta)^s v\,\mathrm{d}y = \int_{\mathbb{S}^N} u\mathcal{P}_{2s}u\,\mathrm{d}y = \sum_{l,m}\lambda_l u_{l,m}^2.$$

Since $C_0^\infty(\mathbb{R}^N)$ is dense in the space $\mathcal{D}^{s,2}(\mathbb{R}^N) := \{v \in L^{\frac{2N}{N-2s}}(\mathbb{R}^N) : (-\Delta)^{s/2}v \in L^2(\mathbb{R}^N)\}$ (see e.g. [7]), the equality

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2}v|^2\,\mathrm{d}y = \sum_{l,m}^{2}\lambda_l u_{l,m}^2 \tag{3.31}$$

extends to all $v \in \mathcal{D}^{s,2}(\mathbb{R}^N)$. In particular it holds for $v \in \widetilde{H}^s(\Omega)$.

Proof of Proposition 3.4. We first prove (3.23) for $(x, \lambda) = (0, 1)$. Let $v \in T_{0,1}^{\perp}$ and denote $u = v_{\mathcal{S}}$. We claim that the orthogonality conditions on v imply that $u_{l,m} = 0$ for $l = 0, 1$. Indeed, e.g. from $v \perp \partial_\lambda PU_{0,1}$ and recalling $J_{\mathcal{S}}(y) = (\frac{2}{1+|y|^2})^N = 2^N U_{0,1}^{\frac{2N}{N-2s}}$ we compute

$$0 = \int_{\mathbb{R}^N} (-\Delta)^{s/2} v (-\Delta)^{s/2} \partial_\lambda PU_{0,1} \, \mathrm{d}y = c_{N,s} \frac{N+2s}{N-2s} \int_{\mathbb{R}^N} v U_{0,1}^{\frac{4s}{N-2s}} \partial_\lambda U_{0,1} \, \mathrm{d}y$$

$$= c_{N,s} \frac{N+2s}{N-2s} 2^{-\frac{N+2s}{2}} \int_{\mathbb{R}^N} v(y) J_{\mathcal{S}}(y)^{\frac{N+2s}{2N}} \frac{1-|y|^2}{1+|y|^2} \, \mathrm{d}y = c_{N,s} \frac{N+2s}{N-2s} 2^{-\frac{N+2s}{2}} \int_{\mathbb{S}^N} u(\omega) \omega_{N+1} \, \mathrm{d}\sigma(\omega).$$

Analogous calculations show that $v \perp PU_{0,1}$ implies $\int_{\mathbb{S}^N} u = 0$ and that $v \perp \partial_{x_i} PU_{0,1}$ implies $\int_{\mathbb{S}^N} u\omega_i = 0$ for $i = 1, ..., N$. Since the functions 1 and ω_i $(i = 1, ..., N+1)$ form a basis of the space of spherical harmonics of angular momenta $l = 0$ and $l = 1$ respectively, we have proved our claim.

Since the eigenvalues λ_l of $\mathcal{P}_{2s}$ are increasing in l, changing back variables to $\mathbb{R}^N$, we deduce from (3.31) that

$$\int_{\mathbb{R}^N} |(-\Delta)^{s/2} v|^2 \, \mathrm{d}y = \sum_{l,m} \lambda_l u_{l,m}^2 \geq \lambda_2 \int_{\mathbb{S}^N} u(\omega)^2 \, \mathrm{d}\sigma(\omega) = 2^{2s} \lambda_2 \int_{\mathbb{R}^N} v^2(y) U_{0,1}^{p-2}(y) \, \mathrm{d}y.$$

By an explicit computation using the numerical values of λ_2 given by (3.29) and $c_{N,s}$ given in Lemma B.5, this is equivalent to

$$\|(-\Delta)^{s/2} v\|_2^2 - c_{N,s}(p-1) \int_{\Omega} U_{0,1}^{p-2} v^2 \, \mathrm{d}y \geq \frac{4s}{N+2s+2} \|(-\Delta)^{s/2} v\|_2^2, \tag{3.32}$$

which is the desired inequality.

The case of general $(x, \lambda) \in \Omega \times (0, \infty)$ can be deduced from this by scaling. Indeed, for $v \in T_{x,\lambda}^{\perp}$, set $v_{x,\lambda}(y) := v(x - \lambda^{-1} y)$. Then $v_{x,\lambda} \in T_{0,1}^{\perp}$ with respect to the set $\lambda(x - \Omega)$, so that by the above $v_{x,\lambda}$ satisfies

$$\|(-\Delta)^{s/2} v_{x,\lambda}\|_2^2 - c_{N,s}(p-1) \int_{\lambda(x-\Omega)} U_{0,1}^{p-2} v_{x,\lambda}^2 \, \mathrm{d}y \geq \frac{4s}{N+2s+2} \|(-\Delta)^{s/2} v_{x,\lambda}\|_2^2.$$

Changing back variables now yields (3.32). $\qquad\square$

3.4. **Improved a priori bounds.** The main section of this section is the following proposition, which improves Proposition 3.3. It states that the concentration point x_0 does not lie on the boundary of Ω and gives an optimal quantitative bound on w.

Proposition 3.6. *As $\varepsilon \to 0$,*

$$\|(-\Delta)^{s/2} w\| = \mathcal{O}(\lambda^{-\frac{N-2s}{2}}) \tag{3.33}$$

and

$$d^{-1} = \mathcal{O}(1). \tag{3.34}$$

In particular, $x_0 \in \Omega$.

The proposition will readily follow from the following expansion of $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$ with respect to the decomposition $u_\varepsilon = \alpha(PU_{x,\lambda} + w)$ obtained in the previous section.

Lemma 3.7. *As $\varepsilon \to 0$, we have*

$$\mathcal{S}_{a+\varepsilon V}[u_\varepsilon] = S + 2^{2s}\pi^{N/2}\frac{\Gamma(s)}{\Gamma(\frac{N-2s}{2})}\left(\frac{S}{c_{N,s}}\right)^{-\frac{N-2s}{2s}}\phi_0(x)\lambda^{-N+2s}$$

$$+ \left(\frac{S}{c_{N,s}}\right)^{-\frac{N-2s}{2s}}\left(\|(-\Delta)^{s/2}w\|^2 + \int_\Omega aw^2\,\mathrm{d}y - c_{N,s}(p-1)\int_\Omega U_{x,\lambda}^{p-2}w^2\,\mathrm{d}y\right)$$

$$+ \mathcal{O}\left(\lambda^{-\frac{N-2s}{2}}\|(-\Delta)^{s/2}w\|\right) + o\left((d\lambda)^{-N+2s}\right) + o\left(\|(-\Delta)^{s/2}w\|^2\right).$$

Proof of Proposition 3.6. Using the almost minimality assumption (1.17) and the coercivity inequality from Proposition 3.5, the expansion from Lemma 3.7 yields the inequality

$$0 \geq (1+o(1))(S-S(a+\varepsilon V))+c\phi_0(x)\lambda^{-N+2s}+c\|(-\Delta)^{s/2}w\|^2+\mathcal{O}\left(\lambda^{-\frac{N-2s}{2}}\|(-\Delta)^{s/2}w\|\right)+o\left((d\lambda)^{-N+2s}\right)$$

for some $c > 0$. By Lemma A.1, we have the lower bound $\phi_0(x) \gtrsim d^{-N+2s}$. Together with the estimate

$$\mathcal{O}\left(\lambda^{-\frac{N-2s}{2}}\|(-\Delta)^{s/2}w\|\right) \leq \delta\|(-\Delta)^{s/2}w\|^2 + C_\delta\lambda^{-N+2s},$$

we obtain by taking δ small enough

$$C_\delta\lambda^{-N+2s} \geq (1+o(1))(S - S(a+\varepsilon V)) + c(d\lambda)^{-N+2s} + c\|(-\Delta)^{s/2}w\|^2.$$

Since all three terms on the right side are nonnegative, the proposition follows. □

Proof of Lemma 3.7. **Step 1:** *Expansion of the numerator.* By orthogonality, we have

$$\|(-\Delta)^{s/2}(PU_{x,\lambda} + w)\|^2 = \|(-\Delta)^{s/2}PU_{x,\lambda}\|^2 + \|(-\Delta)^{s/2}w\|^2.$$

The main term can be written as

$$\|(-\Delta)^{s/2}PU_{x,\lambda}\|^2 = \int_\Omega PU_{x,\lambda}(-\Delta)^s PU_{x,\lambda}\,\mathrm{d}y = c_{N,s}\int_\Omega U_{x,\lambda}^{p-1}PU_{x,\lambda}\,\mathrm{d}y$$

$$= c_{N,s}\int_\Omega U_{x,\lambda}^p\,\mathrm{d}y + c_{N,s}\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-1}H_0(x,\cdot)\,\mathrm{d}y + \mathcal{O}(\|f_{x,\lambda}\|_\infty\int_\Omega U_{x,\lambda}^{p-1}\,\mathrm{d}y),$$

where we used $PU_{x,\lambda} = U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_0(x,\cdot) + f_{x,\lambda}$ with $\|f_{x,\lambda}\|_\infty \lesssim \lambda^{\frac{-N+4-2s}{2}}d^{-N-2+2s}$, by Lemma A.2. Thus

$$\|f\|_\infty\int_\Omega U_{x,\lambda}^{p-1}\,\mathrm{d}y \lesssim (d\lambda)^{-N+2s-2} = o((d\lambda)^{-N+2s}).$$

Next, we have

$$\int_{\mathbb{R}^N\setminus\Omega} U_{x,\lambda}^p\,\mathrm{d}y \leq \int_{\mathbb{R}^N\setminus B_d} U_{x,\lambda}^p\,\mathrm{d}y \lesssim \int_{d\lambda}^\infty \frac{r^{N-1}}{(1+r^2)^N}\,\mathrm{d}r \lesssim (d\lambda)^{-N} = o((d\lambda)^{-N+2s})$$

and thus

$$c_{N,s}\int_\Omega U_{x,\lambda}^p\,\mathrm{d}y = c_{N,s}\|U_{0,1}\|_p^p + o((d\lambda)^{-N+2s}).$$

Finally, using that $H_0(x,y) = \phi_0(x) + \mathcal{O}(\|\nabla_y H_0(x,\cdot)\|_\infty|x-y|) = \phi_0(x) + \mathcal{O}(d^{-N+2s-1}|x-y|)$ by Lemma A.1, we have

$$\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-1}H_0(x,\cdot)\,\mathrm{d}y = \lambda^{-\frac{N-2s}{2}}\phi_0(x)\int_{B_d} U_{x,\lambda}^{p-1}\,\mathrm{d}y + \mathcal{O}(d^{-N+2s-1}\lambda^{-\frac{N-2s}{2}}\int_{B_d} U_{x,\lambda}^{p-1}|x-y|\,\mathrm{d}y)$$

$$+ \lambda^{-\frac{N-2s}{2}}\int_{\Omega\setminus B_d} U_{x,\lambda}^{p-1}H_0(x,y)\,\mathrm{d}y.$$

Since $H_0(x,y) \lesssim d^{-N+2s}$ by Lemma A.1, the last term is

$$\lambda^{-\frac{N-2s}{2}}\int_{\Omega\setminus B_d} U_{x,\lambda}^{p-1}H_0(x,\cdot)\,\mathrm{d}y \lesssim d^{-N+2s}\lambda^{-\frac{N-2s}{2}}\int_{\mathbb{R}^N\setminus B_d} U_{x,\lambda}^{p-1}\,\mathrm{d}y \lesssim (d\lambda)^{-N} = o((d\lambda)^{-N+2s}).$$

Similarly,

$$\lambda^{-\frac{N-2s}{2}}\phi_0(x)\int_{B_d} U_{x,\lambda}^{p-1}\,\mathrm{d}y = \lambda^{-\frac{N-2s}{2}}\phi_0(x)\|U_{0,1}\|_{p-1}^{p-1} + \mathcal{O}\left(\phi_0(x)\lambda^{-\frac{N-2s}{2}}\int_{\mathbb{R}^N\setminus B_d} U_{x,\lambda}^{p-1}\,\mathrm{d}y\right)$$

$$= \lambda^{-\frac{N-2s}{2}}\phi_0(x)\|U_{0,1}\|_{p-1}^{p-1} + \mathcal{O}((d\lambda)^{-N})$$

$$= \lambda^{-\frac{N-2s}{2}}\phi_0(x)\|U_{0,1}\|_{p-1}^{p-1} + o((d\lambda)^{-N+2s}).$$

Finally,

$$d^{-N+2s-1}\lambda^{-\frac{N-2s}{2}}\int_{B_d} U_{x,\lambda}^{p-1}|x-y| \lesssim (d\lambda)^{-N+2s-1}\int_0^{d\lambda}\frac{r^N}{(1+r^2)^{\frac{N+2s}{2}}}\,\mathrm{d}r = o((d\lambda)^{-N+2s})$$

(where one needs to distinguish the cases where $1-2s$ is positive, negative or zero because the dr-integral is divergent if $1-2s \geq 0$).

Collecting all the previous estimates, we have proved

$$\|(-\Delta)^{s/2}PU_{x,\lambda}\|^2 = c_{N,s}\|U_{x,\lambda}\|_p^p + c_{N,s}\lambda^{-N+2s}\|U_{x,\lambda}\|_{p-1}^{p-1}\phi_0(x) + o((d\lambda)^{-N+2s}). \qquad (3.35)$$

The potential term splits as

$$\int_\Omega (a+\varepsilon V)(PU_{x,\lambda}+w)^2\,\mathrm{d}y = \int_\Omega (a+\varepsilon V)PU_{x,\lambda}^2\,\mathrm{d}y + \int_\Omega aw^2\,\mathrm{d}y + \int_\Omega\left((a+\varepsilon V)PU_{x,\lambda}w + \varepsilon Vw^2\right)\mathrm{d}y$$

and we can estimate

$$\left|\int_\Omega (a+\varepsilon V)PU_{x,\lambda}^2\,\mathrm{d}y\right| \lesssim \|U_{x,\lambda}\|_2^2 \lesssim \lambda^{-N+2s}$$

as well as

$$\int_\Omega\left((a+\varepsilon V)PU_{x,\lambda}w + \varepsilon Vw^2\right)\mathrm{d}y \lesssim \|PU_{x,\lambda}\|_{p'}\|w\|_p + \varepsilon\|w\|^2 = \mathcal{O}\left(\lambda^{-\frac{N-2s}{2}}\|(-\Delta)^{s/2}w\|\right) + o(\|(-\Delta)^{s/2}w\|^2).$$

In summary we have, for the numerator of $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$,

$$\alpha^{-2}\left(\|(-\Delta)^{s/2}u\|^2 + \int_\Omega (a+\varepsilon V)u^2\,\mathrm{d}y\right)$$

$$= c_{N,s}\|U_{x,\lambda}\|_p^p + c_{N,s}\lambda^{-N+2s}\|U_{x,\lambda}\|_{p-1}^{p-1}\phi_0(x) + \|(-\Delta)^{s/2}w\|^2 + \int_\Omega aw^2\,\mathrm{d}y$$

$$+ \mathcal{O}(\lambda^{-N+2s}) + \mathcal{O}(\lambda^{-\frac{N-2s}{2}}\|(-\Delta)^{s/2}w\|) + o((d\lambda)^{-N+2s}) + o(\|(-\Delta)^{s/2}w\|^2).$$

Step 2: *Expansion of the denominator.* By Taylor's formula,

$$(PU_{x,\lambda}+w)^p = PU_{x,\lambda}^p + pPU_{x,\lambda}^{p-1}w + \frac{p(p-1)}{2}PU_{x,\lambda}^{p-2}w^2 + \mathcal{O}(PU_{x,\lambda}^{p-3}|w|^3 + |w|^p).$$

Note that, strictly speaking, we use this formula if $p \geq 3$. If $2 < p \leq 3$, the same is true without the remainder term $PU^{p-3}|w|^3$, which does not affect the rest of the proof. To evaluate the main term, write $PU_{x,\lambda} = U_{x,\lambda} - \varphi_{x,\lambda}$ with $\varphi_{x,\lambda} := \lambda^{-1/2}H_0(x,\cdot) + f_{x,\lambda}$, see Lemma A.2. Then

$$\int_\Omega PU_{x,\lambda}^p\,\mathrm{d}y = \int_\Omega U_{x,\lambda}^p\,\mathrm{d}y - p\int_\Omega U_{x,\lambda}^{p-1}\varphi_{x,\lambda}\,\mathrm{d}y + \mathcal{O}\left(\int_\Omega (U_{x,\lambda}^{p-2}\varphi_{x,\lambda}^2 + \varphi_{x,\lambda}^p)\,\mathrm{d}y\right)$$

$$\|U_{0,1}\|_p^p - p\lambda^{-N+2s}\|U_{0,1}\|_{p-1}^{p-1}\phi_0(x) + o((d\lambda)^{-N+2s})$$

where we used that by Lemmas A.2 and B.1 $\int_\Omega U_{x,\lambda}^{p-2}\varphi_{x,\lambda}^2\,\mathrm{d}y \leq \|U_{x,\lambda}\|_{p-2}^{p-2}\|\varphi_{x,\lambda}\|_\infty^2 \lesssim (d\lambda)^{-2N+4s} = o((d\lambda)^{-N+2s})$ and $\|\varphi_{x,\lambda}\|_p^p \lesssim (d\lambda)^{-N} = o((d\lambda)^{-N+2s})$.

Next, the integral of the remainder term is controlled by

$$\int_\Omega (PU_{x,\lambda}^{p-3}|w|^3 + |w|^p)\,\mathrm{d}y \lesssim \|PU_{x,\lambda}\|_p^{p-3}\|w\|_p^3 + \|w\|_p^p = o(\|(-\Delta)^{s/2}w\|^2).$$

The term linear in w is

$$\int_\Omega PU_{x,\lambda}^{p-1} w \, \mathrm{d}y = \int_\Omega U_{x,\lambda}^{p-1} w \, \mathrm{d}y + \mathcal{O}\left(\int_\Omega (U_{x,\lambda}^{p-2}\varphi|w| + \varphi^{p-1}|w|) \, \mathrm{d}y\right).$$

Now by orthogonality of w, we have

$$\int_\Omega U_{x,\lambda}^{p-1} w \, \mathrm{d}y = c_{N,s}^{-1} \int_\Omega (-\Delta)^s U_{x,\lambda} w \, \mathrm{d}y = \int_{\mathbb{R}^N} (-\Delta)^{s/2} U_{x,\lambda} (-\Delta)^{s/2} w \, \mathrm{d}y = 0.$$

Moreover, using $\|\varphi_{x,\lambda}\|_p \lesssim (d\lambda)^{-\frac{N-2s}{2}}$ by Lemma A.2, we get

$$\left|\int_\Omega \varphi_{x,\lambda}^{p-1} w \, \mathrm{d}y\right| \le \|\varphi_{x,\lambda}\|_p^{p-1}\|w\|_p \le (d\lambda)^{-\frac{N+2s}{2}}\|(-\Delta)^{s/2}w\| = o((d\lambda)^{-N+2s}).$$

Using additionally that $\|\varphi_{x,\lambda}\|_\infty \lesssim d^{-N+2s}\lambda^{-\frac{N-2s}{2}}$ by Lemma A.2, by the same computation as in [22, Lemma A.1] we get $\|U_{x,\lambda}^{p-2}\varphi_{x,\lambda}\|_{\frac{p}{p-1}} \lesssim (d\lambda)^{-N+2s}$ and therefore

$$\int_\Omega U_{x,\lambda}^{p-2}\varphi_{x,\lambda}|w| \, \mathrm{d}y \lesssim (d\lambda)^{-N+2s}\|(-\Delta)^{s/2}w\|.$$

In summary we have, for the denominator of $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$,

$$\alpha^{-p}\int_\Omega u_\varepsilon^p \, \mathrm{d}y = \|U_{0,1}\|_p^p - p\lambda^{-N+2s}\|U_{0,1}\|_{p-1}^{p-1}\phi_0(x) + \frac{p(p-1)}{2}\int_\Omega U_{x,\lambda}^{p-2}w^2 \, \mathrm{d}y$$
$$+ \mathcal{O}((d\lambda)^{-N+2s}\|(-\Delta)^{s/2}w\|) + o(\|(-\Delta)^{s/2}w\|^2) + o((d\lambda)^{-N+2s}).$$

Step 3: *Expansion of the quotient.* Using Taylor's formula, we find, for the denominator,

$$\alpha^{-2}\left(\int_\Omega u_\varepsilon^p \, \mathrm{d}y\right)^{-2/p} = \|U_{0,1}\|_p^{-2} + 2\|U_{0,1}\|_p^{-p-2}\|U_{0,1}\|_{p-1}^{p-1}\lambda^{-N+2s}\phi_0(x)$$
$$- c_{N,s}(p-1)\|U_{0,1}\|_p^{-p-2}\int_\Omega U_{x,\lambda}^{p-2}w^2 \, \mathrm{d}y$$
$$+ \mathcal{O}((d\lambda)^{-N+2s}\|(-\Delta)^{s/2}w\|) + o(\|(-\Delta)^{s/2}w\|^2) + o((d\lambda)^{-N+2s}).$$

Multiplying this with the expansion for the denominator found above, we obtain

$$\mathcal{S}_{a+\varepsilon V}[u_\varepsilon] = c_{N,s}\|U_{0,1}\|_p^{p-2} + \lambda^{-N+2s}c_{N,s}\|U_{0,1}\|_p^{-2}\|U_{0,1}\|_{p-1}^{p-1}\phi_0(x)$$
$$\|U_{0,1}\|_p^{-2}\left(\|(-\Delta)^{s/2}w\|^2 + \int_\Omega aw^2 \, \mathrm{d}y - c_{N,s}(p-1)\int_\Omega U_{x,\lambda}^{p-2}w^2 \, \mathrm{d}y\right)$$
$$+ \mathcal{O}((d\lambda)^{-N+2s}\|(-\Delta)^{s/2}w\|) + o(\|(-\Delta)^{s/2}w\|^2) + o((d\lambda)^{-N+2s}).$$

Expressing the various constants using Lemma B.5, we find

$$c_{N,s}\|U_{0,1}\|_p^{p-2} = S,$$

$$\|U_{0,1}\|_p^{-2} = \left(\frac{S}{c_{N,s}}\right)^{-\frac{N-2s}{2s}},$$

$$c_{N,s}\|U_{0,1}\|_p^{-2}\|U_{0,1}\|_{p-1}^{p-1} = 2^{2s}\pi^{N/2}\frac{\Gamma(s)}{\Gamma(\frac{N-2s}{2})}\left(\frac{S}{c_{N,s}}\right)^{-\frac{N-2s}{2s}}.$$

This yields the expansion claimed in the lemma. $\qquad\square$

4. Proof of Theorem 1.2

At this point, we have collected sufficiently precise information on the behavior of a general almost minimizing sequence to prove Theorem 1.2.

The main difficulty of the argument consists in constructing, for a critical potential a, a point $x_0 \in \Omega$ at which $\phi_a(x_0) = 0$. To do so, we carry out some additional analysis for a sequence u_ε which we assume to consist of true minimizers of $S(a - \varepsilon)$, not only almost minimizers as in the rest of this paper. We make this additional assumption essentially for convenience and brevity of the argument, see the remark below Lemma 4.1.

Indeed, since a is critical, we have $S(a - \varepsilon) < S$ for every $\varepsilon > 0$. By the results of [38], which adapts the classical lemma of Lieb contained in [9] to the fractional case, this strict inequality implies existence of a minimizer u_ε of $S(a - \varepsilon)$. Normalizing $\int_\Omega u_\varepsilon^{\frac{2N}{N-2s}} \, \mathrm{d}y = A_{N,s}$ as in (1.17), u_ε satisfies the equation

$$(-\Delta)^s u_\varepsilon + (a - \varepsilon)u_\varepsilon = \frac{S(a - \varepsilon)}{A_{N,s}^{2s/N}} u_\varepsilon^{\frac{N+2s}{N-2s}} \quad \text{on } \Omega, \qquad u \equiv 0 \quad \text{on } \mathbb{R}^N \setminus \Omega. \tag{4.1}$$

By using equation (4.1), we can conveniently extract the leading term of the remainder term w_ε. We do this in the following lemma, which is the key step in the proof of Theorem 1.2.

Lemma 4.1. *Let u_ε be minimizers of $S(a - \varepsilon)$ which satisfy (4.1). Then we have*

$$S(a - \varepsilon) = S + c_{N,s} A_{N,s}^{-2/p} \phi_a(x) \lambda^{-N+2s} + o(\lambda^{-N+2s}). \tag{4.2}$$

If $8s/3 < N$, Lemma 4.1 is in fact implied by the more refined analysis carried out in Section 5 below, which does not use the equation (4.1). If $2s < N \leq 8s/3$, we speculate than one can prove Lemma 4.1 for almost minimizers not satisfying (4.1) by arguing like in [24, Section 5], but we do not pursue this explicitly here.

Proof of Lemma 4.1. Clearly, the analysis carried out in Section 3 so far applies to the sequence (u_ε). Thus, up to passing to a subsequence, we may assume that $u_\varepsilon = \alpha_\varepsilon(PU_{x_\varepsilon, \lambda_\varepsilon} + w_\varepsilon)$ with $\alpha_\varepsilon \to 1$, $x_\varepsilon \to x_0 \in \Omega$ and $\|(-\Delta)^{s/2} w_\varepsilon\|_2 \lesssim \lambda^{-\frac{N-2s}{2}}$ as $\varepsilon \to 0$.

Thus the sequence $\tilde{w}_\varepsilon := \lambda_\varepsilon^{\frac{N-2s}{2}} w_\varepsilon$ is bounded in $\widetilde{H}^s(\Omega)$ and converges weakly in $\widetilde{H}^s(\Omega)$, up to a subsequence, to some $\tilde{w}_0 \in \widetilde{H}^s(\Omega)$. Inserting the expansion $u = \alpha(PU_{x,\lambda} + w)$ in (4.1), the equation fulfilled by $\tilde{w}_\varepsilon$ reads

$$(-\Delta)^s \tilde{w}_\varepsilon + (a - \varepsilon)\tilde{w}_\varepsilon = -(a - \varepsilon)PU_{x,\lambda} \lambda^{\frac{N-2s}{2}} + \lambda^{-2s} \frac{S(a - \varepsilon)}{A_{N,s}^{2s/N}} \left(PU_{x,\lambda} \lambda^{\frac{N-2s}{2}} + \tilde{w}_\varepsilon \right)^{\frac{N+2s}{N-2s}}. \tag{4.3}$$

By Lemma A.2, we can write

$$PU_{x,\lambda} \lambda^{\frac{N-2s}{2}} = G_0(x, \cdot) - \lambda^{N-2s} h(\lambda(x - \cdot)) - \lambda^{N-2s} f_{x,\lambda}$$

with h as in Lemma B.3. By the bounds on h and $f_{x,\lambda}$ from Lemmas A.2 and B.3, this yields

$$PU_{x,\lambda} \lambda^{\frac{N-2s}{2}} \to G_0(x_0, \cdot) \text{ uniformly on compacts of } \Omega \setminus \{x_0\} \text{ and in } L^q(\Omega), q < \tfrac{N}{N-2s}. \tag{4.4}$$

Letting $\varepsilon \to 0$ in (4.3), we obtain

$$\int_\Omega (-\Delta)^{s/2} \tilde{w}_0 (-\Delta)^{s/2} \varphi \, \mathrm{d}y + \int_\Omega a \tilde{w}_0 \varphi \, \mathrm{d}y = -\int_\Omega a G_0(x, \cdot) \varphi \, \mathrm{d}y \tag{4.5}$$

for every $\varphi \in C_c^\infty(\Omega \setminus \{x_0\})$. Now it is straightforward to show that $C_c^\infty(\Omega \setminus \{x_0\})$ is dense in $\widetilde{H}^s(\Omega)$, by using a cutoff function argument together with identity (3.1). Thus by approximation, (4.5) even holds for every $\varphi \in \widetilde{H}^s(\Omega)$. In other words, $\tilde{w}_0$ weakly solves the equation

$$(-\Delta)^s \tilde{w}_0 + a\tilde{w}_0 = -aG_0(x_0, \cdot) \quad \text{on } \Omega, \qquad \tilde{w}_0 \equiv 0 \quad \text{on } \mathbb{R}^N \setminus \Omega.$$

By uniqueness of solutions, we conclude $\tilde{w}_0 = H_0(x_0, \cdot) - H_a(x_0, \cdot)$.

We will now use this information to prove the desired expansion (4.2) of the energy $S(a - \varepsilon) = \mathcal{S}_{a-\varepsilon}[PU_{x,\lambda} + w]$. Indeed, using the already established bound $\|(-\Delta)^{s/2}w\| \lesssim \lambda^{-\frac{N-2s}{2}}$, the numerator is

$$\|(-\Delta)^{s/2}PU_{x,\lambda}\|^2 + \int_\Omega a(PU_{x,\lambda}^2 + 2PU_{x,\lambda}w)\,\mathrm{d}y + \|(-\Delta)^{s/2}w\|^2 + \int_\Omega aw^2\,\mathrm{d}y + o(\lambda^{-N+2s}). \quad (4.6)$$

By integrating the equation for w against w and recalling $\frac{S(a-\varepsilon)}{A_{N,s}^{2s/N}} = c_{N,s} + o(1)$, we easily find the asymptotic identity (compare [21, eq. (8)] for $s = 1$)

$$\|(-\Delta)^{s/2}w\|^2 + \int_\Omega aw^2\,\mathrm{d}y = c_{N,s}(p-1)\int_\Omega U_{x,\lambda}^{p-2}w^2\,\mathrm{d}y - \int_\Omega aPU_{x,\lambda}w\,\mathrm{d}y.$$

Inserting this in (4.6), together with the expansion of $\|(-\Delta)^{s/2}PU_{x,\lambda}\|^2$ given in (3.35), the numerator of $\mathcal{S}_{a-\varepsilon}[PU_{x,\lambda} + w]$ becomes

$$c_{N,s}A_{N,s} - c_{N,s}a_{N,s}\phi_0(x)\lambda^{-N+2s} + \int_\Omega a(PU_{x,\lambda}^2 + PU_{x,\lambda}w)\,\mathrm{d}y + c_{N,s}(p-1)\int_\Omega U_{x,\lambda}^{p-2}w^2\,\mathrm{d}y + o(\lambda^{-N+2s}).$$

$$(4.7)$$

The numerator of $\mathcal{S}_{a-\varepsilon}[PU_{x,\lambda} + w]$, by the computations in the proof of Lemma 3.7, is given by

$$\left(\int_\Omega (PU_{x,\lambda} + w)^p\,\mathrm{d}y\right)^{-2/p} = A_{N,s}^{-\frac{2}{p}} - A_{N,s}^{-\frac{2}{p}-1}\left(-2\phi_0(x)\lambda^{-N+2s} + (p-1)\int_\Omega U_{x,\lambda}^{p-2}w^2\,\mathrm{d}y\right). \quad (4.8)$$

Multiplying out (4.7) and (4.8), the terms in $\int_\Omega U^{p-2}w^2\,\mathrm{d}y$ cancel precisely and we obtain

$$S(a - \varepsilon) = S + c_{N,s}A_{N,s}^{-2/p}a_{N,s}\phi_0(x)\lambda^{-N+2s} \tag{4.9}$$

$$+ A_{N,s}^{-2/p}\lambda^{-N+2s}\left(\int_\Omega a\left((\lambda^{\frac{N-2s}{2}}PU_{x,\lambda})^2 + \lambda^{\frac{N-2s}{2}}PU_{x,\lambda}\tilde{w}\right)\,\mathrm{d}y\right) + o(\lambda^{-N+2s}).$$

Now we are ready to return to our findings about $\tilde{w}_0$. Indeed, by (4.4), and observing that $G_0(x, \cdot)$ is an admissible test function in (4.5), we get

$$\int_\Omega a\left((\lambda^{\frac{N-2s}{2}}PU_{x,\lambda})^2 + \lambda^{\frac{N-2s}{2}}PU_{x,\lambda}\tilde{w}\right)\,\mathrm{d}y = \int_\Omega a\left(G_0(x, \cdot)^2 + G_0(x, \cdot)\tilde{w}_0\right)\,\mathrm{d}y + o(1) \tag{4.10}$$

$$= -\int_\Omega (-\Delta)^{s/2}\tilde{w}_0(-\Delta)^{s/2}G_0(x, \cdot)\,\mathrm{d}y + o(1) = -\tilde{w}_0(x) = \gamma_{N,s}(\phi_a(x) - \phi_0(x)) + o(1). \tag{4.11}$$

By inserting this into (4.9) and observing that $\gamma_{N,s} = c_{N,s}a_{N,s}$ by the numerical values given in Lemma B.5, the proof is complete. $\qquad \square$

Now we have all the ingredients to give a quick proof of our first main result.

Proof of Theorem 1.2. As explained after the statement of the theorem, it only remains to prove the implication $(ii) \Rightarrow (i)$. Suppose thus $S(a) < S$ and let $c > 0$ be the smallest number such that $\bar{a} := a + c$ satisfies $S(\bar{a}) = S$. For $\varepsilon > 0$, let u_ε be the sequence of minimizers $S(\bar{a} - \varepsilon)$, normalized to satisfy (4.1). By Lemma 4.1, we have

$$S > S(\bar{a} - \varepsilon) = S - c_{N,s}A_{N,s}^{-2/p}\phi_a(x)\lambda^{-N+2s} + o(\lambda^{-N+2s}).$$

Letting $\varepsilon \to 0$, this shows $\phi_{\bar{a}}(x_0) \leq 0$. By the resolvent identity, we have for every $x \in \Omega$

$$\phi_{\tilde{a}}(x) = \phi_a(x) + \int_\Omega (\tilde{a} - a)(z) G_a(x, z) G_{a+c}(z, x) \, \mathrm{d}z < \phi_a(x),$$

and hence $\phi_a(x_0)$ is strictly monotone in a. Thus $\phi_a(x_0) < \phi_{a+c}(x_0) = 0$, and the proof is complete. □

5. Proof of the lower bound II: a refined expansion

This section is the most technical of the paper. It is devoted to extracting the leading term of the remainder w and to obtaining sufficiently good bounds on the new error term. In Section 5.2 we will need to work under the additional assumption $8s/3 < N$ in order to obtain the required precision.

Concretely, we write

$$w = \lambda^{-\frac{N-2s}{2}} (H_0(x, \cdot) - H_a(x, \cdot)) + q$$

and decompose the remainder further into a tangential and an orthogonal part

$$q = t + r, \qquad t \in T_{x,\lambda}, \quad r \in T_{x,\lambda}^\perp.$$

(We keep omitting the subscript ε.) A refined expansion of $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$ then yields an error term in r which can be controlled using the coercivity inequality of Proposition 3.5. The refined expansion is derived in Section 5.2 below.

On the other hand, since t is an element of the $(N + 2)$-dimensional space $T_{x,\lambda}$, it can be bounded by essentially explicit computations. This is achieved in Section 5.1.

Remark 5.1. The present Section 5 thus constitutes the analogon of [24, Section 6], where the same analysis is carried out for the case $s = 1$ and $N = 3$. We emphasize that, despite these similarities, our approach is conceptually somewhat simpler than that of [24]. Indeed, the argument in [24] relies on an intermediate step involving a spectral cutoff construction, through which the apriori bound $\|\nabla q\| = o(\lambda^{1/2}) = o(\lambda^{-\frac{N-2s}{2}})$ is obtained.

On the contrary, we are able to conduct the following analysis with only the weaker bound $\|\nabla q\| = \mathcal{O}(\lambda^{-\frac{N-2s}{2}})$ at hand (which follows from Proposition 3.6). This comes at the price of some additional explicit error terms in r, which can however be conveniently absorbed (see Lemmas 5.7 and 5.9). Since $N > 8s/3$ is fulfilled when $N = 3$, $s = 1$, this simplified proof of course also works in the particular situation of [24].

5.1. **A precise description of t.** For λ large enough, the functions $PU_{x,\lambda}$, $\partial_\lambda PU_{x,\lambda}$ and $\partial_{x_i} PU_{x,\lambda}$, $i = 1, ..., N$ are linearly independent. There are therefore uniquely determined coefficients β, γ, δ_i, $i = 1, ..., N$, such that

$$t = \beta \lambda^{-N+2s} PU_{x,\lambda} + \gamma \lambda^{-N+2s+1} \partial_\lambda PU_{x,\lambda} + \sum_{i=1}^N \delta_i \lambda^{-N+2s-2} \partial_{x_i} PU_{x,\lambda}. \tag{5.1}$$

Here the choice of the different powers of λ multiplying the coefficients is justified by the following result.

Lemma 5.2. *As $\varepsilon \to 0$, we have $\beta, \gamma, \delta_i = \mathcal{O}(1)$.*

As a corollary, we obtain estimates on t in various norms.

Lemma 5.3. *As $\varepsilon \to 0$,*

$$\|(-\Delta)^{s/2} t\|_2 \lesssim \lambda^{-N+2s} \quad and \quad \|t\|_{\frac{2N}{N+2s}} \lesssim \|t\|_2 \lesssim \lambda^{-\frac{3N-6s}{2}}.$$

Proof. Recall that $PU_{x,\lambda} = U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}} H_0(x,\cdot) + f_{x,\lambda}$. Then all bounds follow in a straightforward way from (5.1) together with Lemma 5.2 and the standard bounds from Lemmas A.1, A.2, B.1 and B.2. $\qquad\square$

Proof of Lemma 5.2. **Step 1.** We introduce the normalized basis functions

$$\tilde{\varphi}_1 := \frac{PU_{x,\lambda}}{\|(-\Delta)^{s/2}PU_{x,\lambda}\|}, \quad \tilde{\varphi}_2 := \frac{\partial_\lambda PU_{x,\lambda}}{\|(-\Delta)^{s/2}\partial_\lambda PU_{x,\lambda}\|}, \quad \tilde{\varphi}_j := \frac{\partial_{x_{j-2}}PU_{x,\lambda}}{\|(-\Delta)^{s/2}\partial_{x_{j-2}}PU_{x,\lambda}\|}, \tag{5.2}$$

and prove that

$$a_j := \int_{\mathbb{R}^N} (-\Delta)^{s/2}\tilde{\varphi}_j (-\Delta)^{s/2} t \, \mathrm{d}y = \begin{cases} \mathcal{O}(\lambda^{-N+2s}), & j = 1,2, \\ \mathcal{O}(\lambda^{-N+2s-1}), & j = 3,...,N+2. \end{cases} \tag{5.3}$$

Since $\lambda^{-\frac{N-2s}{2}}(H_0(x,\cdot) - H_a(x,\cdot)) + t + r = w \in T_{x,\lambda}^\perp$, and $r \in T_{x,\lambda}^\perp$, we have

$$a_j = \lambda^{-\frac{N-2s}{2}} \int_{\mathbb{R}^N} (-\Delta)^{s/2}\tilde{\varphi}_j (-\Delta)^{s/2} (H_a(x,\cdot) - H_0(x,\cdot)) \, \mathrm{d}y$$

$$= \lambda^{-\frac{N-2s}{2}} \int_{\mathbb{R}^N} (-\Delta)^s \tilde{\varphi}_j (H_a(x,\cdot) - H_0(x,\cdot)) \, \mathrm{d}y.$$

Thus,

$$a_1 = \lambda^{-\frac{N-2s}{2}} \|(-\Delta)^{s/2}PU_{x,\lambda}\|^{-1} c_{N,s} \int_\Omega U_{x,\lambda}^{\frac{N+2s}{N-2s}} (H_a(x,\cdot) - H_0(x,\cdot)) \, \mathrm{d}y$$

$$\lesssim \lambda^{-N+2s},$$

where we used that by Lemma B.2, $\|(-\Delta)^{s/2}PU_{x,\lambda}\|^{-1} \lesssim 1$. The bound for a_2 follows similarly. To obtain the claimed improved bound for a_j, $j = 3,...,N+2$, we write

$$a_{i+2} \lesssim \lambda^{-\frac{N-2s}{2}} \|(-\Delta)^{s/2}\partial_{x_i}PU_{x,\lambda}\|^{-1} \int_\Omega U_{x,\lambda}^{\frac{4s}{N-2s}} \partial_{x_i}U (H_a(x,\cdot) - H_0(x,\cdot)) \, \mathrm{d}y$$

$$\lambda^{-\frac{N-2s}{2}} \|(-\Delta)^{s/2}\partial_{x_i}PU_{x,\lambda}\|^{-1} \mathcal{O}\left(\int_{\mathbb{R}^N \setminus B_d} U_{x,\lambda}^{\frac{4s}{N-2s}} |\partial_{x_i}U_{x,\lambda}| \, \mathrm{d}y + \int_\Omega U_{x,\lambda}^{\frac{4s}{N-2s}} |\partial_{x_i}U_{x,\lambda}||x-y| \, \mathrm{d}y \right)$$

$$\lesssim \lambda^{-N+2s-1}.$$

Here we wrote $H_a(x,y) - H_0(x,y) = \phi_a(x) - \phi_0(x) + \mathcal{O}(|x-y|)$ and used that by oddness of $\partial_{x_i}U$,

$$(\phi_a(x) - \phi_0(x)) \int_{B_d} U_{x,\lambda}^{\frac{4s}{N-2s}} \partial_{x_i}U_{x,\lambda} = 0.$$

This concludes the proof of (5.3).

Step 2. We write

$$t = \sum_{j=1}^{N+2} b_j \tilde{\varphi}_j,$$

with

$$b_1 = \beta\lambda^{-N+2s}\|(-\Delta)^{s/2}PU_{x,\lambda}\|, \quad b_2 = \gamma\|(-\Delta)^{s/2}\partial_\lambda PU_{x,\lambda}\|,$$

$$b_j = \delta_j\lambda^{-N+2s-2}\|(-\Delta)^{s/2}\partial_{x_{j-2}}PU_{x,\lambda}\|, \quad j = 3,...,N+2.$$

Our goal is to show that

$$b_j = a_j + \mathcal{O}(\lambda^{-N+2s}) \sup_k a_k, \quad j = 1,..,N+2. \tag{5.4}$$

From (5.4) we conclude by the estimates on the a_j from (5.3) and Lemma B.2.

To prove (5.4), we define the Gram matrix G by

$$G_{j,k} := ((-\Delta)^{s/2}\tilde{\varphi}_j, (-\Delta)^{s/2}\tilde{\varphi}_k).$$

By Lemma B.2 and the definition of the $\tilde{\varphi}_j$ it is easily checked that

$$G_{j,k} = \delta_{j,k} + \mathcal{O}(\lambda^{-N+2s}), \qquad j, k = 1, ..., N + 2.$$

Thus for sufficiently large λ, G is invertible with

$$(G^{-1})_{j,k} = \delta_{j,k} + \mathcal{O}(\lambda^{-N+2s}). \tag{5.5}$$

By definition of G,

$$\psi_j := \sum_{k=1}^{N+2} (G^{-1/2})_{j,k}\tilde{\varphi}_k \tag{5.6}$$

is an orthonormal basis of $T_{x,\lambda}$. We can therefore write

$$
\begin{aligned}
t &= \sum_j \left((-\Delta)^{s/2}\psi_j, (-\Delta)^{s/2}t \right) \psi_j = \sum_j \sum_k (G^{-1/2})_{j,k} \left((-\Delta)^{s/2}\tilde{\varphi}_k, (-\Delta)^{s/2}t \right) \phi_j \\
&= \sum_j \sum_k (G^{-1/2})_{j,k} a_k \sum_l (G^{-1/2})_{j,l}\tilde{\varphi}_l \\
&= \sum_l \left(\sum_k \left(\sum_j (G^{-1/2})_{l,j}(G^{-1/2})_{j,k} \right) a_k \right) \tilde{\varphi}_l \\
&= \sum_l \left(\sum_k (G^{-1})_{l,k} a_k \right) \tilde{\varphi}_l
\end{aligned}
$$

Thus $b_l = \sum_k (G^{-1})_{l,k} a_k$ and (5.4) follows from (5.5). $\qquad \square$

Remark 5.4. By treating the terms in the above proof more carefully, it can be shown in fact that $\lambda^{N-2s}\beta$, $\lambda^{N-2s-1}\gamma$ and $\lambda^{N-2s+2}\delta_i$ have a limit as $\lambda \to \infty$. Indeed, for instance, the leading orders of the expressions $\int_\Omega U_{x,\lambda}^{\frac{N+2s}{N-2s}}(H_a(x, \cdot) - H_0(x, \cdot)) \, \mathrm{d}y$ and $\|(-\Delta)^{s/2}PU_{x,\lambda}\|$ going into the leading behavior of β can be explicitly evaluated, see Lemma A.4 and the proof of Lemma B.2 respectively. We do not need the behavior of the coefficients β, γ, δ_i to that precision in the following, so we do not state them explicitly.

5.2. **The new expansion of $\mathcal{S}_{a+\varepsilon V}[u]$.** Our goal is now to expand the value of the energy functional $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$ with respect to the refined decomposition introduced above, namely

$$u = \alpha(\psi_{x,\lambda} + q) = \alpha \left(PU_{x,\lambda} + \left(\lambda^{-\frac{N-2s}{2}} H_a(x, \cdot) - H_0(x, \cdot) \right) + t + r \right).$$

In all that follows, we work under the important assumption that

$$-3N + 6s < -2s, \quad \text{i.e. } \tfrac{8}{3}s < N \tag{5.7}$$

so that $\lambda^{-3N+6s} = o(\lambda^{-2s})$. Assumption (5.7) has the consequence that, using the available bounds on t and r, we can expand the energy $\mathcal{S}_{a+\varepsilon V}[u]$ up to $o(\lambda^{-2s})$ errors in a way that does not depend on t. This is the content of the next lemma.

Lemma 5.5. *As $\varepsilon \to 0$, we have*

$$\mathcal{S}_{a+\varepsilon V}[u_\varepsilon] = \mathcal{S}_{a+\varepsilon V}[\psi_{x,\lambda}] + D_0^{-2/p} \left(\mathcal{E}_0[r] - \frac{2N_0}{pD_0} I[r] + o(\|(-\Delta)^{s/2}r\|^2) \right)$$

$$+ o(\lambda^{-2s}) + o(\varepsilon\lambda^{-N+2s}) + o(\phi_a(x)\lambda^{-N+2s}).$$

Here,

$$N_0 := \|(-\Delta)^{s/2}\psi_{x,\lambda}\|_2^2 + \int_\Omega (a + \varepsilon V)\psi_{x,\lambda}^2 \, \mathrm{d}y, \qquad D_0 := \int_\Omega \psi_{x,\lambda}^p \, \mathrm{d}y, \tag{5.8}$$

and $I[r]$ is as defined in (5.10) below.

We emphasize that the contribution of t enters only into the remainders $o(\lambda^{-2s}) + o(\varepsilon\lambda^{-N+2s}) + o(\phi_a(x)\lambda^{-N+2s})$. This is remarkable because t enters to orders $\lambda^{-N+2s} >> \lambda^{-2s}$ and $\lambda^{-2N+4s} >> \lambda^{-2s}$ (if $N < 3s$) into both the numerator and the denominator of $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$, see Lemmas 5.6 and 5.7 below. When calculating the quotients, these contributions cancel precisely, as we verify in Lemma 5.8 below. Heuristically, such a phenomenon is to be expected because (up to projection onto $\widetilde{H}^s(\Omega)$ and perturbation by $a + \varepsilon V$) by definition t represents the directions along which the quotient functional is invariant. As already pointed out in the introduction, we suspect, but cannot prove, that in the absence of assumption (5.7) the contributions of t to the higher order coefficients $\lambda^{-kN+2ks}$ for $3 \leq k \leq \frac{2N}{N-2s}$ would continue to cancel.

We prove Lemma 5.5 by separately expanding the numerator and the denominator of $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$. We abbreviate

$$\mathcal{E}_\varepsilon[u] := \|(-\Delta)^{s/2}u\|^2 + \int_\Omega (a + \varepsilon V)u^2 \, \mathrm{d}y$$

and write $\mathcal{E}_\varepsilon[u, v]$ for the associated bilinear form.

Lemma 5.6 (Expanding the numerator). *As $\varepsilon \to 0$,*

$$|\alpha|^{-2}\mathcal{E}_\varepsilon[u_\varepsilon] = \mathcal{E}_\varepsilon[\psi_{x,\lambda}] + \left(2\mathcal{E}_0[\psi_{x,\lambda}, t] + \|(-\Delta)^{s/2}t\|^2\right) + \mathcal{E}_0[r] + o(\lambda^{-2s}) + o(\varepsilon\lambda^{-N+2s}).$$

Proof. We write $\alpha^{-1}u_\varepsilon = \psi_{x,\lambda} + t + r$ and therefore

$$\mathcal{E}_\varepsilon[u_\varepsilon] = \mathcal{E}_\varepsilon[\psi_{x,\lambda}] + 2\mathcal{E}_\varepsilon[\psi_{x,\lambda}, t + r] + \mathcal{E}_\varepsilon[t + r]. \tag{5.9}$$

The third term on the right side is

$$\mathcal{E}_\varepsilon[t + r] = \mathcal{E}_0[t] + 2\mathcal{E}_0[t, r] + \mathcal{E}_0[r] + \varepsilon\int_\Omega V(t + r)^2 \, \mathrm{d}y.$$

Now $\int_{\mathbb{R}^N}(-\Delta)^{s/2}t(-\Delta)^{s/2}r \, \mathrm{d}y = 0$ by orthogonality and therefore, by Lemma 5.3,

$$\mathcal{E}_0[t, r] = \int_\Omega atr \, \mathrm{d}y = \mathcal{O}(\|(-\Delta)^{s/2}r\|\|t\|_{\frac{2N}{N+2s}}) = \mathcal{O}(\lambda^{\frac{-3N+6s}{2}}\|(-\Delta)^{s/2}r\|) = o(\lambda^{-2s}) + o(\|(-\Delta)^{s/2}r\|^2),$$

where the last equality is a consequence of assumption (5.7) and Young's inequality. Finally, again by Lemma 5.3,

$$\varepsilon\int_\Omega V(t + r)^2 \, \mathrm{d}y = \mathcal{O}(\varepsilon(\|t\|^2 + \|(-\Delta)^{s/2}r\|^2)) = o(\varepsilon\lambda^{-N+2s}) + o(\|(-\Delta)^{s/2}r\|^2).$$

The second term on the right side of (5.9) is

$$2\mathcal{E}_\varepsilon[\psi_{x,\lambda}, t + r] = 2\mathcal{E}_0[\psi_{x,\lambda}, t] + 2\mathcal{E}_0[\psi_{x,\lambda}, r] + 2\varepsilon\int_\Omega V\psi_{x,\lambda}(t + r) \, \mathrm{d}y.$$

To start with, using Lemma 5.3,

$$\varepsilon\int_\Omega V\psi_{x,\lambda}(t + r) \, \mathrm{d}y = \mathcal{O}(\varepsilon(\|t\|_{\frac{2N}{N+2s}} + \|\psi\|_{\frac{2N}{N+2s}}\|(-\Delta)^{s/2}r\|))$$
$$= \mathcal{O}(\varepsilon\lambda^{-\frac{N-2s}{2}}\|(-\Delta)^{s/2}r\|) + o(\varepsilon\lambda^{-N+2s}) = o(\varepsilon\lambda^{-N+2s}) + o(\|(-\Delta)^{s/2}r\|^2)),$$

again by Young's inequality. Moreover, using that $r \in T_{x,\lambda}^\perp$, that $(-\Delta)^s H_a(x, \cdot) = aG_a(x, \cdot)$ and $(-\Delta)^s H_0(x, \cdot) = 0$, and integrating by parts,

$$\mathcal{E}_0[\psi_{x,\lambda}, r] = \lambda^{-\frac{N-2s}{2}}\int_\Omega (-\Delta)^{s/2}(H_0 - H_a)(x, \cdot)(-\Delta)^{s/2}r \, \mathrm{d}y + \int_\Omega a\psi_{x,\lambda}r \, \mathrm{d}y$$
$$= \int_\Omega a(-\lambda^{-\frac{N-2s}{2}}G_a(x, \cdot) + \psi_{x,\lambda})r \, \mathrm{d}y.$$

Since we can write $\lambda^{-\frac{N-2s}{2}} G_a(x, y) - \psi_{x,\lambda}(y) = \lambda^{\frac{N-2s}{2}} h(\lambda(x - y)) + f_{x,\lambda}$ with h as in Lemma B.3 and $f_{x,\lambda}$ as in Lemma A.2, we get

$$\mathcal{E}_0[\psi_{x,\lambda}, r] \lesssim \left(\lambda^{-\frac{N-2s}{2}} \|h(\lambda \cdot)\|_{\frac{2N}{N+2s}} + \|f_{x,\lambda}\|_\infty\right) \|(-\Delta)^{s/2} r\| \lesssim \lambda^{-2s} \|(-\Delta)^{s/2} r\| = o(\lambda^{-2s})$$

by the bounds in those lemmas. Finally,

$$\mathcal{E}_0[t] = \|(-\Delta)^{s/2} t\|^2 + \int_\Omega at^2 \, dy$$

and $\int_\Omega at^2 \, dy \lesssim \|t\|_2^2 \lesssim \lambda^{-3N+6s} = o(\lambda^{-2s})$ by Lemma 5.3 and Assumption (5.7). $\qquad\square$

Lemma 5.7 (Expanding the denominator)**.** *As $\varepsilon \to 0$,*

$$|\alpha|^{-p} \int_\Omega u_\varepsilon^p \, dy = \int_\Omega \psi_{x,\lambda}^p \, dy + \left(p \int_\Omega \psi_{x,\lambda}^{p-1} t \, dy + \frac{p(p-1)}{2} \int_\Omega \psi_{x,\lambda}^{p-2} t^2 \, dy\right) + \frac{p(p-1)}{2} \int_\Omega \psi_{x,\lambda}^{p-2} r^2 \, dy$$
$$+ \mathcal{O}\left(\lambda^{-\frac{N-2s}{2}} \int_\Omega U^{p-2} |H_a| |r| \, dy\right) + o(\|(-\Delta)^{s/2} r\|^2) + o(\lambda^{-2s}).$$

Proof. Write $\alpha^{-1} u_\varepsilon = \psi_{x,\lambda} + t + r$. We expand

$$\int_\Omega (\psi_{x,\lambda} + t + r)^p \, dy = \int_\Omega (\psi_{x,\lambda} + r)^p \, dy + p \int_\Omega (\psi_{x,\lambda} + r)^{p-1} t \, dy + \frac{p(p-1)}{2} \int_\Omega (\psi_{x,\lambda} + r)^{p-2} t^2 \, dy$$
$$+ \mathcal{O}\left(\|\psi_{x,\lambda} + r\|_p^{p-3} \|(-\Delta)^{s/2} t\|^3 + \|(-\Delta)^{s/2} t\|^p\right).$$

By Lemma 5.3 together with assumption (5.7), the last term is $o(\lambda^{-2s})$. The third term is, by Lemma 5.3,

$$\int_\Omega (\psi_{x,\lambda} + r)^{p-2} t^2 \, dy = \int_\Omega \psi_{x,\lambda}^{p-2} t^2 \, dy + \mathcal{O}\left(\lambda^{-2N+4s} \|(-\Delta)^{s/2} r\|\right).$$

The second term is

$$\int_\Omega (\psi_{x,\lambda} + r)^{p-1} t \, dy = \int_\Omega \psi_{x,\lambda}^{p-1} t \, dy + (p-1) \int_\Omega \psi_{x,\lambda}^{p-2} rt \, dy + o(\|(-\Delta)^{s/2} r\|^2).$$

The remaining term $\int_\Omega \psi_{x,\lambda}^{p-2} rt \, dy$ needs to be expanded more carefully. Using $\psi_{x,\lambda} = U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}} H_a(x, \cdot) - f_{x,\lambda}$ with $\|\lambda^{-\frac{N-2s}{2}} H_a(x, \cdot) + f_{x,\lambda}\|_\infty \lesssim \lambda^{-\frac{N-2s}{2}}$, we write

$$\int_\Omega \psi_{x,\lambda}^{p-2} rt \, dy = \int_\Omega U_{x,\lambda}^{p-2} rt \, dy + \mathcal{O}\left(\lambda^{-\frac{N-2s}{2}} \|(-\Delta)^{s/2} r\| \|(-\Delta)^{s/2} t\|\right)$$

and using assumption (5.7), the remainder is bounded by

$$\lambda^{-\frac{N-2s}{2}} \|(-\Delta)^{s/2} r\| \|(-\Delta)^{s/2} t\| \lesssim \lambda^{\frac{-3N+6s}{2}} \|(-\Delta)^{s/2} r\| = o(\lambda^{-2s}) + o(\|(-\Delta)^{s/2} r\|^2).$$

Now using orthogonality of r and the expansion (5.1) of s, by some standard calculations, whose details we omit, one obtains

$$\int_\Omega U_{x,\lambda}^{p-2} rt \, dy = \mathcal{O}\left(\lambda^{\frac{-3N+6s}{2}} \right\| (-\Delta)^{s/2} r\| = o(\lambda^{-2s}) + o(\|(-\Delta)^{s/2} r\|^2),$$

where we used again assumption (5.7) for the last equality.

It remains only to treat the t-independent term $\int_\Omega (\psi_{x,\lambda} + r)^p \, dy$. We find

$$\int_\Omega (\psi_{x,\lambda} + r)^p \, dy = \int_\Omega \psi_{x,\lambda}^p \, dy + p \int_\Omega \psi_{x,\lambda}^{p-1} r \, dy + \frac{p(p-1)}{2} \int_\Omega \psi_{x,\lambda}^{p-2} r^2 \, dy + o(\|(-\Delta)^{s/2} r\|^2).$$

Using orthogonality of r, we get that $\int_\Omega U_{x,\lambda}^{p-1} r \, dy = 0$ and hence

$$\int_\Omega \psi_{x,\lambda}^{p-1} r \, dy = \mathcal{O}\left(\lambda^{-\frac{N-2s}{2}} \int_\Omega U_{x,\lambda}^{p-2} |H_a(x, \cdot)| |r| \, dy + \lambda^{-\frac{N+2s}{2}} \|(-\Delta)^{s/2} r\|\right)$$

and $\lambda^{-\frac{N+2s}{2}}\|(-\Delta)^{s/2}r\| \lesssim \lambda^{-N} = o(\lambda^{-2s})$. Finally, we have

$$\int_\Omega \psi_{x,\lambda}^{p-2} r^2 \, \mathrm{d}y = \int_\Omega U_{x,\lambda}^{p-2} r^2 \, \mathrm{d}y + o(\|(-\Delta)^{s/2}r\|^2).$$

Collecting all the estimates gives the claim of the lemma.   □

We can now prove the claimed expansion of the energy functional.

Proof of Lemma 5.5. We write the expansions of the numerator and the denominator as

$$\mathcal{E}_\varepsilon[u_\varepsilon] = N_0 + N_1 + \mathcal{E}_0(r) + o(\lambda^{-2s} + (\varepsilon + \phi_a(x))\lambda^{-N+2s}),$$

where

$$N_0 = \mathcal{E}_\varepsilon[\psi], \quad N_1 := 2\mathcal{E}_0[\psi_{x,\lambda}, t] + \|(-\Delta)^{s/2}t\|^2,$$

and

$$\int_\Omega u_\varepsilon^p = D_0 + D_1 + \mathcal{I}[r] + o(\lambda^{-2s}),$$

where

$$D_0 = \int_\Omega \psi^p, \qquad D_1 := p\int_\Omega \psi_{x,\lambda}^{p-1} t + \frac{p(p-1)}{2}\int_\Omega \psi_{x,\lambda}^{p-2} t^2,$$

and

$$\mathcal{I}[r] := \frac{p(p-1)}{2}\int_\Omega \psi_{x,\lambda}^{p-2} r^2 + \mathcal{O}(\lambda^{-\frac{N-2s}{2}}\int_\Omega U^{p-2}|H_a||r| \, \mathrm{d}y). \tag{5.10}$$

Taylor expanding up to and including second order, we find

$$\left(\int_\Omega u_\varepsilon^p\right)^{-2/p} = D_0^{-2/p}\left(1 - \frac{2}{p}\frac{D_1 + I[r]}{D_0} + \frac{p+2}{p^2}\frac{(D_1 + I[r])^2}{D_0^2}\right) + o(\lambda^{-2s}).$$

We now observe $\mathcal{I}[r] \lesssim \|(-\Delta)^{s/2}r\|^2 + o(\phi_a(x)\lambda^{-N+2s} + \lambda^{-2s})$ since

$$\lambda^{-\frac{N-2s}{2}}\int_\Omega U^{p-2}|H_a||r| \lesssim \int_\Omega U^{p-2}r^2 + \lambda^{-N+2s}\int_\Omega U^{p-2}H_a^2 \lesssim \|(-\Delta)^{s/2}r\|^2 + o(\lambda^{-N+2s}\phi_a(x))$$

Hence we can simplify the expression of the denominator to

$$\left(\int_\Omega u_\varepsilon^p\right)^{-2/p} = D_0^{-2/p}\left(1 - \frac{2}{p}\frac{D_1 + I[r]}{D_0} + \frac{p+2}{p^2}\frac{D_1^2}{D_0^2}\right) + o(\|(-\Delta)^{s/2}r\|^2) + o(\phi_a(x)\lambda^{-N+2s}) + o(\lambda^{-2s}).$$

Multiplying this with the expansion of the numerator from above, we find

$$\mathcal{S}_{a+\varepsilon V}[u_\varepsilon] = D_0^{-2/p}N_0 + D_0^{-2/p}\left(N_1 - \frac{2}{p}\frac{N_0}{D_0}D_1 - \frac{2}{p}\frac{N_1 D_1}{D_0} + \frac{p+2}{p^2}\frac{D_1^2 N_0}{D_0^2}\right)$$

$$+ D_0^{-2/p}\left(\mathcal{E}_0[r] - \frac{2N_0}{pD_0}I[r] + o(\|(-\Delta)^{s/2}r\|^2)\right) + o(\lambda^{-2s}) + o(\phi_a(x)\lambda^{-N+2s}).$$

We show in Lemma 5.8 below that the bracket involving the terms N_1 and D_1 involving s vanishes up to order $o(\lambda^{-2s})$, due to cancellations. Noting that $D_0^{-2/p}N_0$ is nothing but $\mathcal{S}_{a+\varepsilon V}[\psi]$, the expansion claimed in Lemma 5.5 follows.   □

Lemma 5.8. *Assume* (5.7) *and let* N_0, N_1, D_0, D_1 *be defined as in the proof of Lemma 5.5. Then*

$$N_1 = 2\beta c_{N,s}A_{N,s}\lambda^{-N+2s}$$

$$+ c_{N,s}\lambda^{-2N+4s}\left(\beta^2 A_{N,s} + \gamma^2(p-1)B_{N,s} - 2a_{N,s}\phi_0(x)(\beta - \frac{N-2s}{2}\gamma)\right) + o(\lambda^{-2s})$$

and

$$D_1 = \lambda^{-N+2s} p\beta A_{N,s} + \lambda^{-2N+4s}\left(\frac{p(p-1)}{2}(\beta^2 A_{N,s} + \gamma^2 B_{N,s}) - p(\beta - \frac{N-2s}{2}\gamma)a_{N,s}\phi_0(x)\right)$$
$$+ o(\phi_a(x)\lambda^{-N+2s}) + o(\lambda^{-2s}).$$

where we abbreviated $B_{N,s} := \int_{\mathbb{R}^N} U_{0,1}^{p-2}|\partial_\lambda U_{0,1}|^2 \, \mathrm{d}y$.

In particular,

$$N_1 - \frac{2}{p}\frac{N_0}{D_0}D_1 - \frac{2}{p}\frac{N_1 D_1}{D_0} + \frac{p+2}{p^2}\frac{D_1^2 N_0}{D_0^2} = o(\lambda^{-2s}) + o(\phi_a(x)\lambda^{-N+2s}).$$

Proof. We start with expanding $N_1 = 2\mathcal{E}_0[\psi, t] + \|(-\Delta)^{s/2}t\|^2$. From Lemma B.2 and the expansion (5.1) for t, one easily sees that

$$\|(-\Delta)^{s/2}t\|^2 = \beta^2\lambda^{-2N+4s}\|(-\Delta)^{s/2}PU_{x,\lambda}\|^2 + \gamma^2\lambda^{-2N+4s+2}\|(-\Delta)^{s/2}\partial_\lambda PU_{x,\lambda}\|^2$$
$$= \beta^2 c_{N,s}A_{N,s}\lambda^{-2N+4s} + \gamma^2(p-1)c_{N,s}B_{N,s} + o(\lambda^{-2s}),$$

where we also used assumption (5.7). Next, recalling $((-\Delta)^s + a)\psi_{x,\lambda} = c_{N,s}U_{x,\lambda}^{p-1} - a(\lambda^{\frac{N-2s}{2}}h(\lambda(x - \cdot)) + f_{x,\lambda})$ with h as in Lemma B.3, we easily obtain

$$2\mathcal{E}_0[\psi_{x,\lambda}, t] = 2c_{N,s}\int_\Omega U_{x,\lambda}^{p-1}t\,\mathrm{d}y + o(\lambda^{-2s}) = 2\beta c_{N,s}A_{N,s}\lambda^{-N+2s}$$
$$- 2c_{N,s}a_{N,s}\phi_0(x)\lambda^{-2N+4s}(\beta + \frac{N-2s}{2}\gamma) + o(\lambda^{-2s}).$$

(Observe that the leading order term with γ vanishes because $\int_{\mathbb{R}^N} U_{0,1}^{p-1}\partial_\lambda U_{0,1} = 0$.) This proves the claimed expansion for N_1. For D_1, we have

$$\int_\Omega \psi_{x,\lambda}^{p-1}t\,\mathrm{d}y = \lambda^{-N+2s}\beta\int_\Omega \psi_{x,\lambda}^{p-1}PU_{x,\lambda}\,\mathrm{d}y + \gamma\lambda^{-N+2s+1}\int_\Omega \psi_{x,\lambda}^{p-1}\partial_\lambda PU_{x,\lambda}\,\mathrm{d}y + o(\lambda^{-2s}).$$

Writing out $\psi_{x,\lambda} = U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_0(x, \cdot) - f_{x,\lambda}$ and $PU_{x,\lambda} = U_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_0(x, \cdot) - f$, by the usual bounds together with assumption (5.7) we get

$$\lambda^{-N+2s}\beta\int_\Omega \psi^{p-1}PU_{x,\lambda}\,\mathrm{d}y = \lambda^{-N+2s}\beta A_{N,s} - \lambda^{-2N+4s}\beta a_{N,s}\phi_0(x) + o(\lambda^{-N+2s}\phi_a(x)) + o(\lambda^{-2s}).$$

Similarly,

$$\gamma\lambda^{-N+2s+1}\int_\Omega \psi^{p-1}\partial_\lambda PU_{x,\lambda}\,\mathrm{d}y = \gamma\frac{N-2s}{2}\lambda^{-2N+4s}a_{N,s}\phi_0(x) + o(\lambda^{-N+2s}\phi_a(x)) + o(\lambda^{-2s}).$$

(Observe that the leading order term with γ vanishes because $\int_{\mathbb{R}^N} U_{0,1}^{p-1}\partial_\lambda U_{0,1} = 0$.) Finally,

$$\int_\Omega \psi_{x,\lambda}^{p-2}t^2\,\mathrm{d}y = \lambda^{-2N+4s}\left(\beta^2 A_{N,s} + \gamma^2 B_{N,s}\right) + o(\lambda^{-2s}).$$

Putting together the above, we end up with the claimed expansion for D_1.

The last assertion of the lemma follows from the expansions of N_0, D_0, N_1 and D_1 by an explicit calculation whose details we omit. $\qquad\square$

Based on the refined expansion of $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$ obtained in Lemma 5.5, we are now in a position to give the proofs of our main results.

We first use the coercivity inequality from Proposition 3.4 to control the terms involving r that appear in Lemma 5.5.

Lemma 5.9 (Coercivity result). *There is $\rho > 0$ such that, as $\varepsilon \to 0$,*

$$\mathcal{E}_0[r] - \frac{2N_0}{pD_0}I[r] \geq \rho\|(-\Delta)^{s/2}r\|^2.$$

Proof. Recalling the definition (5.10) of $\mathcal{I}[r]$ and observing that $N_0/D_0 = c_{N,s}$, we find by Proposition 3.5 that

$$\mathcal{E}_0[r] - \frac{2N_0}{pD_0}I[r] = \|(-\Delta)^{s/2}r\|^2 + \int_\Omega ar^2\,\mathrm{d}y - c_{N,s}(p-1)\int_\Omega U_{x,\lambda}^{p-2}r^2\,\mathrm{d}y$$

$$+ \mathcal{O}\left(\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-2}|H_a(x\cdot)||r|\,\mathrm{d}y\right) + o(\|(-\Delta)^{s/2}r\|^2)$$

$$\geq \rho\|(-\Delta)^{s/2}r\|^2 + \mathcal{O}\left(\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-2}|H_a(x\cdot)||r|\,\mathrm{d}y\right)$$

for some $\rho > 0$. The remaining error term can be bounded as follows.

$$\lambda^{-\frac{N-2s}{2}}\int_\Omega U_{x,\lambda}^{p-2}|H_a(x,\cdot)||r|\,\mathrm{d}y \leq \delta'\int_\Omega U_{x,\lambda}^{p-2}r^2\,\mathrm{d}y + C\lambda^{-N+2s}\int_\Omega U_{x,\lambda}^{p-2}H_a(x,\cdot)^2\,\mathrm{d}y$$

$$\leq \delta\|(-\Delta)^{s/2}r\|^2 + \mathcal{O}(\lambda^{-2N+4s}\phi_a(x)^2) + o(\lambda^{-2s})$$

$$\leq \delta\|(-\Delta)^{s/2}r\|^2 + o(\lambda^{-N+2s}\phi_a(x) + \lambda^{-2s})$$

where we used Lemma A.4. By choosing $\delta > 0$ small enough, we obtain the conclusion. $\qquad\square$

6. PROOF OF THE MAIN RESULTS

Combining Lemma 5.9 with Lemma 5.5 gives a lower bound on $\mathcal{S}_{a+\varepsilon V}[u_\varepsilon]$. Using the almost-minimizing assumption (1.17) and the expansion from Theorem 2.1, this lower bound can be stated as follows:

$$0 \geq (1 + o(1))(S - S(a + \varepsilon V)) + \mathcal{R} \tag{6.1}$$

$$+ A_{N,s}^{-\frac{N-2s}{N}}\left((Q_V(x) + o(1))\varepsilon\lambda^{-N+2s} - (a(x) + o(1))(\alpha_{N,s} + c_{N,s}d_{N,s}b_{N,s})\lambda^{-2s}\right),$$

where

$$\mathcal{R} := A_{N,s}^{\frac{N-2s}{N}}\left(a_{N,s}c_{N,s}(1 + o(1))\phi_a(x)\lambda^{-N+2s} + \mathcal{T}_2(\phi_a(x),\lambda)\right) + \rho\|(-\Delta)^{s/2}r\|_2^2,$$

for some $\rho > 0$, and $\mathcal{T}_2(\phi_a(x),\lambda)$ as in (2.5).

Recall that $\phi_a \geq 0$ by Corollary 2.2 and that $\phi_a(x)$ is bounded because $x_0 \in \Omega$. Since $\mathcal{T}_2$ is a sum of higher powers $(\phi_a(x)\lambda^{-N+2s})^k$ with $k \geq 2$, we have $\mathcal{R} \geq 0$ for ε small enough.

Lemma 6.1. *As $\varepsilon \to 0$, $\phi_a(x) = o(1)$. In other words, $x_0 \in \mathcal{N}_a$.*

Proof. Since $S - S(a + \varepsilon V) \geq 0$ and $\|(-\Delta)^{s/2}r\|_2^2 \geq 0$, the bound (6.1) gives

$$\phi_a(x) \lesssim \varepsilon + \lambda^{N-4s} + \lambda^{N-2s}\mathcal{T}_2(\phi_a(x),\lambda).$$

Since $\phi_a(x)$ is uniformly bounded, we can bound $\mathcal{T}_2(\phi_a(x),\lambda) \lesssim \lambda^{-2N+4s}$, which concludes. $\qquad\square$

Lemma 6.2. *If $\mathcal{N}_a(V) \neq \emptyset$, then $x_0 \in \mathcal{N}_a(V)$.*

In the proof of this lemma, we need the assumption (1.12), i.e. that $a(x) < 0$ on $\mathcal{N}_a$.

Proof. By Lemma 6.1 we only need to prove that $Q_V(x_0) < 0$.

Inserting the upper bound from Corollary 2.3 on $S - S(a + \varepsilon V)$ into (6.1), and using $\mathcal{R} \geq 0$, we obtain that

$$(Q_V(x) + o(1))\varepsilon\lambda^{-N+2s} \leq -C_1\varepsilon^{\frac{2s}{4s-N}} + C_2\lambda^{-2s}.$$

Here the numbers C_1 and C_2 are given by

$$C_1 := (1 + o(1))\sigma_{N,s} \sup_{x \in \mathcal{N}_a(V).} \frac{|Q_V(x)|^{\frac{2s}{4s-N}}}{|a(x)|^{\frac{N-2s}{4s-N}}}, \qquad C_2 := -a(x) + o(1)$$

Using Lemma 6.1 and the assumption $a < 0$ on $\mathcal{N}_a$, we have that C_2 is strictly positive and remains bounded away from zero by assumption. Since $\mathcal{N}_a(V)$ is not empty, the same is clearly true for C_1. Thus by Young's inequality

$$-C_1 \varepsilon^{\frac{2s}{4s-N}} + C_2 \lambda^{-2s} \leq -c\varepsilon\lambda^{-N+2s}$$

for some $c > 0$. This implies $Q_V(x_0) < -c < 0$ as desired. $\qquad\square$

Now we are ready to prove our main results Theorems 1.4, 1.5 and 1.6.

Proof of Theorem 1.4. By using $\mathcal{R} \geq 0$ and minimizing the last term over λ, like in the proof of Corollary 2.3, the bound (6.1) implies

$$(1 + o(1))(S(a + \varepsilon V) - S) \geq -\sigma_{N,s} \frac{|Q_V(x_0)|^{\frac{2s}{4s-N}}}{|a(x_0)|^{\frac{N-2s}{4s-N}}} \varepsilon^{\frac{2s}{4s-N}} + o(\varepsilon^{\frac{2s}{4s-N}})$$

$$\geq -\sigma_{N,s} \sup_{x \in \mathcal{N}_a(V)} \frac{|Q_V(x)|^{\frac{2s}{4s-N}}}{|a(x)|^{\frac{N-2s}{4s-N}}} \varepsilon^{\frac{2s}{4s-N}} + o(\varepsilon^{\frac{2s}{4s-N}}),$$

where the last inequality follows from Lemma 6.2. This is equivalent to

$$S(a + \varepsilon V) \geq S - \sigma_{N,s} \sup_{x \in \mathcal{N}_a(V)} \frac{|Q_V(x)|^{\frac{2s}{4s-N}}}{|a(x)|^{\frac{N-2s}{4s-N}}} \varepsilon^{\frac{2s}{4s-N}} + o(\varepsilon^{\frac{2s}{4s-N}}).$$

Since the matching upper bound has already been proved in Corollary 2.3, the proof of the theorem is complete. $\qquad\square$

Proof of Theorem 1.5. Since $x_0 \in \mathcal{N}_a$ by Lemma 6.1, by assumption we have $Q_V(x_0) \geq 0$ and $a(x_0) < 0$. Together with $\mathcal{R} \geq 0$, the bound (6.1) then implies

$$0 \geq (1 + o(1))(S - S(a + \varepsilon V)) + c\lambda^{-2s} + o(\varepsilon\lambda^{-N+2s})$$

for some $c > 0$. Since $o(\varepsilon\lambda^{-N+2s}) \geq -\frac{c}{2}\lambda^{-2s} + o(\varepsilon^{\frac{2s}{4s-N}})$ by Young, this implies $S(a + \varepsilon V) \geq S + o(\varepsilon^{\frac{2s}{N-4s}})$. Since the inequality

$$S(a + \varepsilon V) \leq S \tag{6.2}$$

always holds (e.g. by Theorem 2.1), we obtain $S(a + \varepsilon V) \geq S + o(\varepsilon^{\frac{2s}{N-4s}})$ as desired.

Now assume that additionally $Q_V(x_0) > 0$. With $\mathcal{R} \geq 0$, (6.1) implies, for $\varepsilon > 0$ small enough and some $C_1, C_2 > 0$

$$S(a + \varepsilon V) - S \geq C_1 \varepsilon\lambda^{-N+2s} + C_2 \lambda^{-2s} > 0,$$

which contradicts (6.2). Thus assumption (3.19), under which we have worked so far, cannot be satisfied, and we must have $S(a + \varepsilon_0 V) = S$ for some $\varepsilon_0 > 0$. Since $S(a + \varepsilon V)$ is concave in ε (being the infimum of functions linear in ε) and since $S(a) = S$, we must have $S(a + \varepsilon V) = S$ for all $\varepsilon \in [0, \varepsilon_0]$. $\qquad\square$

Proof of Theorem 1.6. We may first observe that the upper and lower bounds on $S(a + \varepsilon V)$ already discussed in the proof of Theorem 1.4 imply

$$\frac{|Q_V(x_0)|^{\frac{2s}{4s-N}}}{|a(x_0)|^{\frac{N-2s}{4s-N}}} = \sup_{x \in \mathcal{N}_a} \frac{|Q_V(x)|^{\frac{2s}{4s-N}}}{|a(x)|^{\frac{N-2s}{4s-N}}}. \tag{6.3}$$

Now, by using additionally Lemma B.6, the estimate (6.1) becomes

$$(1 + o(1))((S(a + \varepsilon V) - S) \geq -\sigma_{N,s} \frac{|Q_V(x_0)|^{\frac{2s}{4s-N}}}{|a(x_0)|^{\frac{N-2s}{4s-N}}} \varepsilon^{\frac{2s}{4s-N}} + \mathcal{R}' + o(\varepsilon^{\frac{2s}{4s-N}}), \tag{6.4}$$

where

$$
\mathcal{R}' = \begin{cases} \mathcal{R} + c_0 \varepsilon^{\frac{2s-2}{4s-N}} \left(\lambda^{-1} - \lambda_0(\varepsilon)^{-1} \right)^2 & \text{if} \quad \left(\frac{A_\varepsilon}{B_\varepsilon} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}} \lambda^{-1} \leq 2 \left(\frac{2s}{N-2s} \right)^{\frac{1}{N-4s}}, \\ \mathcal{R} + c_0 \varepsilon^{\frac{2s}{4s-N}} & \text{if} \quad \left(\frac{A_\varepsilon}{B_\varepsilon} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}} \lambda^{-1} > 2 \left(\frac{2s}{N-2s} \right)^{\frac{1}{N-4s}}, \end{cases}
$$

in the notation of Lemma B.6, for $A_\varepsilon := A_{N,s}^{-\frac{N-2s}{N}} (\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s})(|a(x_0)| + o(1))$ and $B_\varepsilon := A_{N,s}^{-\frac{N-2s}{N}} (|Q_V(x_0)| + o(1))$, with $\lambda_0(\varepsilon)$ given by (B.5). Now applying in (6.4) the upper bound on $S(a + \varepsilon V)$ from Corollary 2.3 yields

$$
\mathcal{R}' = o(\varepsilon^{\frac{2s}{4s-N}}).
$$

The terms that make up $\mathcal{R}'$ being separately nonnegative, this implies $\mathcal{R} = o(\varepsilon^{\frac{2s}{4s-N}})$ and $(\lambda^{-1} - \lambda_0(\varepsilon)^{-1})^2 = o(\varepsilon^{\frac{2}{4s-N}})$, that is,

$$
\begin{aligned}
\lambda &= \left(\frac{2sA_\varepsilon}{(N-2s)B_\varepsilon} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}} + o(\varepsilon^{-\frac{1}{4s-N}}) \\
&= \left(\frac{2s(\alpha_{N,s} + c_{N,s} d_{N,s} b_{N,s})|a(x_0)|}{(N-2s)|Q_V(x_0)|} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}} + o(\varepsilon^{-\frac{1}{4s-N}})
\end{aligned} \tag{6.5}
$$

and

$$
\|(-\Delta)^{s/2} r\|_2 = o(\varepsilon^{\frac{s}{4s-N}}). \tag{6.6}
$$

Inserting the asymptotics of λ back into $\mathcal{R} = o(\varepsilon^{\frac{2s}{4s-N}})$ now gives

$$
\phi_a(x) = o(\varepsilon). \tag{6.7}
$$

It remains to derive the claimed expansion for α. From Lemma 5.7, we deduce

$$
|\alpha|^{-\frac{2N}{N-2s}} \int_\Omega u_\varepsilon^{\frac{2N}{N-2s}} \, dy = \int_\Omega \psi_{x,\lambda}^{\frac{2N}{N-2s}} \, dy + \mathcal{O} \left(\|(-\Delta)^{s/2} s\|_2 + \|(-\Delta)^{s/2} r\|_2^2 + \lambda^{-N+2s} \right).
$$

Using the bound $\|(-\Delta)^{s/2} s\|_2 \lesssim \lambda^{-N+2s}$, together with (6.5), (6.6) and the expansion of $\int_\Omega \psi_{x,\lambda}^{\frac{2N}{N-2s}}$ from Theorem 2.1, we obtain

$$
|\alpha|^{-p} = 1 + \mathcal{O}(\lambda^{-N+2s}) = 1 + \mathcal{O}(\varepsilon^{\frac{N-2s}{4s-N}}).
$$

This completes the proof of Theorem 1.6. $\qquad\square$

APPENDIX A. GREEN'S FUNCTION

A.1. **The Green's function G_0 and the projections $PU_{x,\lambda}$.** We begin by studying the case $a = 0$. The next lemma collect some important estimates on the regular part $H_0(\cdot, \cdot)$ of the Green's function and the Robin function $\phi_0(x) = H_0(x, x)$, which will turn out very important for our analysis. Similar estimates for $s = 1$ have been derived in [34, Section 2 and Appendix A].

We denote in the following $d(x) := \text{dist}(x, \partial\Omega)$.

Lemma A.1. *Let $x \in \Omega$ and $N > 2s$. Then $y \mapsto H_0(x, y)$ is continuous on Ω and we have, for all $y \in \Omega$,*

$$
0 \leq H_0(x, y) \lesssim d(x)^{2s-N},
$$
$$
|\nabla_y H_0(x, y)| \lesssim d(x)^{2s-N-1}.
$$

Moreover, the Robin function ϕ_0 satisfies the two-sided bound

$$
d(x)^{2s-N} \lesssim \phi_0(x) \lesssim d(x)^{2s-N}. \tag{A.1}
$$

Proof. $H_0(x, \cdot)$ satisfies

$$(-\Delta)^s H_0(x, \cdot) = 0 \qquad \text{on } \Omega,$$

$$H_0(x, \cdot) = \frac{1}{|x - \cdot|^{N-2s}} \qquad \text{on } \mathbb{R}^N \setminus \Omega.$$

Thus we can write

$$H_0(x, y) = \int_{\mathbb{R}^N \setminus \Omega} \frac{1}{|x - z|^{N-2s}} \, dP_\Omega^y(z)$$

where P_Ω^y denotes harmonic measure for $(-\Delta)^s$, see [30, Theorem 7.2]. Since P_Ω^y is a probability measure, this implies

$$0 \leq H_0(x, y) \lesssim d(x)^{-N+2s}.$$

Similarly, since

$$(-\Delta)^s \nabla H_0(x, \cdot) = \nabla (-\Delta)^s H_0(x, \cdot) = 0 \qquad \qquad \text{on } \Omega,$$

$$\nabla H_0(x, \cdot) = \nabla \frac{1}{|x - \cdot|^{N-2s}} \qquad \qquad \text{on } \mathbb{R}^N \setminus \Omega,$$

we have

$$|\nabla_y H_0(x, y)| = \left| \int_{\mathbb{R}^N \setminus \Omega} \left(\nabla_z \frac{1}{|x - z|^{N-2s}} \right) dP_\Omega^y(z) \right| \lesssim d(x)^{-N+2s-1}.$$

The lower bound $d(x)^{2s-N} \lesssim \phi_0(x)$ is proved in [17, Lemma 7.6]. $\qquad \qquad \square$

The following important lemma shows the relation between the regular part $H_0(x, \cdot)$ and the projections $PU_{x,\lambda}$ introduced in (1.16). For the classical case $s = 1$, this is [34, Proposition 1]. For fractional $s \in (0, 1)$, a slightly weaker version relying on the extension formulation of $(-\Delta)^s$ appears in [15, Lemma C.1].

Lemma A.2. *Let $x \in \Omega$ and $N > 2s$.*

(i) *We have*

$$0 \leq PU_{x,\lambda} \leq U_{x,\lambda} \tag{A.2}$$

and the function $\varphi_{x,\lambda} := U_{x,\lambda} - PU_{x,\lambda}$ satisfies the estimates

$$\|\varphi_{x,\lambda}\|_{L^\infty(\mathbb{R}^N)} \lesssim d(x)^{-N+2s} \lambda^{\frac{-N+2s}{2}} \tag{A.3}$$

and

$$\|\varphi_{x,\lambda}\|_{L^p(\mathbb{R}^N)} \lesssim (d(x)\lambda)^{-\frac{N-2s}{2}}. \tag{A.4}$$

(ii) *Moreover, the expansion*

$$PU_{x,\lambda} = U_{x,\lambda} - \lambda^{\frac{N-2s}{2}} H_0(x, y) + f_{x,\lambda}, \tag{A.5}$$

holds with

$$\|f_{x,\lambda}\|_{L^\infty(\Omega)} \lesssim d(x)^{-N-2+2s} \lambda^{-\frac{N+4-2s}{2}}$$

Proof. Claim (i). Our proof follows mostly [34, Appendix A]. Since

$$(-\Delta)^s PU_{x,\lambda} \geq 0 \qquad \text{on } \Omega,$$

$$PU_{x,\lambda} \equiv 0 \qquad \text{on } \mathbb{R}^N \setminus \Omega.$$

the maximum principle (see e.g. [39, Proposition 2.17]) implies that $PU_{x,\lambda} \geq 0$. Similarly, $\varphi_{x,\lambda} = U_{x,\lambda} - PU_{x,\lambda}$ satisfies

$$(-\Delta)^s \varphi_{x,\lambda} = 0 \qquad \qquad \text{on } \Omega, \tag{A.6}$$

$$\varphi_{x,\lambda} = U_{x,\lambda} \geq 0 \qquad \qquad \text{on } \mathbb{R}^N \setminus \Omega.$$

and thus $\varphi_{x,\lambda} \geq 0$ by the maximum principle. This completes the proof of (A.2).

By (A.6), we can moreover write

$$\varphi_{x,\lambda}(y) = \int_{\mathbb{R}^N \setminus \Omega} U_{x,\lambda}(z) \, \mathrm{d}P_\Omega^y(z), \quad y \in \Omega.$$

Thus $\|\varphi_{x,\lambda}\|_{L^\infty(\mathbb{R}^N)} = \|U_{x,\lambda}\|_{L^\infty(\mathbb{R}^N \setminus \Omega)} \lesssim \lambda^{\frac{-N+2s}{2}} d(x)^{-N+2s}$.

Next, let us prove the L^p estimate on $\varphi_{x,\lambda}$. Since $\varphi_{x,\lambda} \in H^s(\mathbb{R}^N)$, by the Sobolev inequality we have

$$\|\varphi_{x,\lambda}\|_{L^p(\mathbb{R}^N)}^2 \lesssim \|(-\Delta)^{s/2}\varphi_{x,\lambda}\|_2^2$$
$$= \|(-\Delta)^{s/2}U_{x,\lambda}\|_2^2 + \|(-\Delta)^{s/2}PU_{x,\lambda}\|_2^2 - 2\int_{\mathbb{R}^N}(-\Delta)^s U_{x,\lambda} PU_{x,\lambda}\,\mathrm{d}y. \tag{A.7}$$

The second summand in (A.7) can be written as

$$\|(-\Delta)^{s/2}PU_{x,\lambda}\|_2^2 = c_{N,s}\int_\Omega PU_{x,\lambda}U_{x,\lambda}^{p-1}\,\mathrm{d}y = \|(-\Delta)^{s/2}U_{x,\lambda}\|_2^2 - c_{N,s}\int_\Omega \varphi_{x,\lambda}U_{x,\lambda}^{p-1}\,\mathrm{d}y$$
$$= \|(-\Delta)^{s/2}U_{x,\lambda}\|_2^2 + \mathcal{O}\left(\|\varphi_{x,\lambda}\|_\infty \int_\Omega U_{x,\lambda}^{p-1}\,\mathrm{d}y\right) = \|(-\Delta)^{s/2}U_{x,\lambda}\|_2^2 + \mathcal{O}\left((d(x)\lambda)^{-N+2s}\right)$$

by (A.3).

Similarly, the third summand in (A.7) is

$$-2\int_{\mathbb{R}^N}(-\Delta)^s U_{x,\lambda}PU_{x,\lambda}\,\mathrm{d}y = -2c_{N,s}\int_\Omega U_{x,\lambda}^p\,\mathrm{d}y + 2c_{N,s}\int_\Omega U_{x,\lambda}^{p-1}\varphi_{x,\lambda}\,\mathrm{d}y$$
$$= -2\|(-\Delta)^{s/2}U_{x,\lambda}\|_2^2 + \mathcal{O}\left((d(x)\lambda)^{-N+2s}\right),$$

where we also used the bound

$$\int_{\mathbb{R}^N \setminus \Omega} U_{x,\lambda}^p\,\mathrm{d}y \lesssim (d(x)\lambda)^{-N}.$$

Collecting these estimates and returning to (A.7), we obtain

$$\|\varphi_{x,\lambda}\|_{L^p(\mathbb{R}^N)}^2 \lesssim (d(x)\lambda)^{-N+2s}.$$

This concludes the proof of (A.4).

Claim (ii). The function $f_{x,\lambda} := \varphi_{x,\lambda} - \lambda^{-\frac{N-2s}{2}}H_0(x,\cdot)$ satisfies

$$(-\Delta)^s f_{x,\lambda} = 0 \qquad\qquad \text{on } \Omega,$$
$$f_{x,\lambda} = U_{x,\lambda} - \frac{\lambda^{-\frac{N-2s}{2}}}{|x-\cdot|^{N-2s}} \qquad\qquad \text{on } \mathbb{R}^N \setminus \Omega.$$

As in the proof of Lemma A.1, we have

$$f_{x,\lambda}(y) = \int_{\mathbb{R}^N \setminus \Omega}\left(U_{x,\lambda}(y) - \frac{\lambda^{-\frac{N-2s}{2}}}{|x-z|^{N-2s}}\right)\mathrm{d}P_\Omega^y(z),$$

and hence, since P_Ω^x is a probability measure, we have

$$\|f_{x,\lambda}\|_{L^\infty(\Omega)} \leq \left\|U_{x,\lambda}(y) - \frac{\lambda^{-\frac{N-2s}{2}}}{|x-y|^{N-2s}}\right\|_{L^\infty(\mathbb{R}^N \setminus \Omega)}$$
$$= \mathcal{O}\left(\lambda^{-\frac{N+4-2s}{2}}\operatorname{dist}(x,\partial\Omega)^{-N-2+2s}\right)$$

by Lemma B.3 below. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

A.2. **Expanding the regular part** $H_b(x,y)$ **near the diagonal.** We now turn to the Green's function G_b, for a general potential $b \in C^1(\Omega) \cap C(\overline{\Omega})$ such that $(-\Delta)^s + b$ is coercive. By noting the potential b rather than a, we emphasize the fact that criticality of b is not needed for the following expansions. Moreover, in contrast to the previous subsection, we specialize to the condition $2s < N < 4s$ again, which plays a role in the proof of Lemmas A.3 and A.4 below.

Lemma A.3. *Let $x \in \Omega$ and $2s < N < 4s$.*

(i) If $4s - N < 1$, then as $y \to x$

$$H_b(x,y) = \phi_b(x) - d_{N,s}b(x)|x-y|^{4s-N} + o(|x-y|^{4s-N}).$$

(ii) If $4s - N \geq 1$, then there is $\xi_x \in \mathbb{R}^N$ such that

$$H_b(x,y) = \phi_b(x) + \xi_x \cdot (y-x) - d_{N,s}b(x)|x-y|^{4s-N} + o(|x-y|^{4s-N}).$$

Here the constant $d_{N,s} > 0$ is given by (B.2). The asymptotics are uniform for x in compact subsets of Ω.

Proof. Fix $x \in \Omega$ and let

$$\psi_x(y) := H_b(x,y) - \phi_b(x) + d_{N,s}b(x)|x-y|^{4s-N},$$

with $d_{N,s}$ as in (B.2). We use the facts that, in the distributional sense,

$$(-\Delta)^s_y H_b(x,y) = b(y)G_b(x,y) = \frac{b(y)}{|x-y|^{N-2s}} - b(y)H_b(x,y)$$

and, by Lemma B.5,

$$(-\Delta)^s|x|^{4s-N} = -d_{N,s}^{-1}|x|^{2s-N}.$$

Thus ψ_x solves, in the distributional sense, the equation

$$(-\Delta)^s_y \psi_x(y) = F_x(y), \tag{A.8}$$

with

$$F_x(y) = \frac{b(y) - b(x)}{|x-y|^{N-2s}} - b(y)H_b(x,y).$$

Since $b \in C^1(\Omega)$, we have

$$\frac{|b(x) - b(y)|}{|x-y|^{N-2s}} \lesssim |x-y|^{-N+2s+1}.$$

We will deduce the assertion of the lemma in each case from elliptic estimates on the equation (A.8) and appropriate bounds on F_x.

Case $-N + 2s + 1 < 0$. Since the second summand $b(y)H_b(x,y)$ is in L^∞, we have $F_x \in L^p(\Omega)$ for every $p < \frac{N}{N-2s-1}$. For the following, fix some $p \in (\frac{N}{2s}, \frac{N}{N-2s-1})$. (The assumption $N < 4s$ guarantees that this interval is not empty.)

Define $\tilde{\psi}_x := (-\Delta)^{-s}F_x$, where $(-\Delta)^{-s}$ is convolution with the Riesz potential. Then by [35, Theorem 1.6.(iii)] we have $[\tilde{\psi}_x]_{C^\alpha(\mathbb{R}^N)} \lesssim \|F_x\|_{L^p(\mathbb{R}^N)}$, where $\alpha = 2s - \frac{N}{p}$. Moreover $(-\Delta)^s(\psi_x - \tilde{\psi}_x) = 0$ on Ω. Since s-harmonic functions are smooth (see e.g. [1, Section 2]), we conclude that $\psi_x \in C^{2s - \frac{N}{p}}(B_{d/2}(x))$.

Since $\psi_x(x) = 0$, we conclude that as $y \to x$,

$$\psi_x(y) = \mathcal{O}(|x-y|^{2s - \frac{N}{p}}). \tag{A.9}$$

If we choose $p \in (\frac{N}{N-2s}, \frac{N}{N-2s-1})$, then $2s - \frac{N}{p} > 4s - N$. (As a consequence of $N < 4s$, we have the inclusion $(\frac{N}{N-2s}, \frac{N}{N-2s-1}) \subset (\frac{N}{2s}, \frac{N}{N-2s-1})$. Together with the definition of ψ_x, (A.9) then implies

$$H_b(x,y) = \phi_b(x) - d_{N,s}b(x)|x-y|^{4s-N} + o(|x-y|^{4s-N}),$$

which is the assertion of the lemma.

Case $-N + 2s + 1 \geq 0$. In this case $F_x \in L^\infty(\Omega)$. More precisely, we have

$$F_x \in \begin{cases} L^\infty(\Omega) & \text{if } N = 2s + 1, \\ C^{0,-N+2s+1}(\Omega) & \text{if } 0 < -N + 2s + 1. \end{cases}$$

Notice that we always have $-N + 2s + 1 < 1$, since $N > 2s$. As above, define $\tilde{\psi}_x = (-\Delta)^{-s} F_x$. By [39], we find using $N < 4s$ that in any of the above cases, $\tilde{\psi}_x \in C^{1,\alpha}$ for all $\alpha \in (0, 1]$ with $\alpha < 4s - N$. Using Hölder continuity of the gradient, we easily find

$$\psi_x(y) = \psi_x(x) + \nabla\psi_x(x) \cdot (y - x) + \mathcal{O}(|x - y|^{\alpha+1}).$$

Choosing $\alpha > 4s - N - 1$ and inserting the definition of ψ_x, we find

$$H_b(x, y) = \phi_b(x) + \nabla\psi_x(x) \cdot (y - x) - d_{N,s}b(x)|x - y|^{4s-N} + o(|x - y|^{4s-N}),$$

which is the assertion of the lemma with $\xi_x := \nabla\psi_x(x)$. $\square$

Lemma A.4. *Let $k \in \mathbb{N}$ with $k \leq p = \frac{2N}{N-2s}$. If $k > \frac{2s}{N-2s}$, then*

$$\lambda^{-\frac{k}{2}(N-2s)} \int_\Omega U_{x,\lambda}^{p-k} H_a(x, \cdot)^k \, dy = o(\lambda^{-2s}).$$

If $2 \leq k \leq \frac{2s}{N-2s}$, then

$$\lambda^{-\frac{k}{2}(N-2s)} \int_\Omega U_{x,\lambda}^{p-k} H_a(x, \cdot)^k \, dy = \left(\int_{\mathbb{R}^N} U_{0,1}(y)^{p-k} \, dy \right) \phi_a(x)^k \lambda^{-k(N-2s)} + o(\lambda^{-2s}).$$

If $k = 1$,

$$\lambda^{-\frac{N-2s}{2}} \int_\Omega U_{x,\lambda}^{p-1} H_b(x, \cdot) \, dy = a_{N,s}\phi_b(x)\lambda^{-N+2s} - d_{N,s}b_{N,s}b(x)\lambda^{-2s} + o(\lambda^{-2s}) + o(\phi_b(x)\lambda^{-N+2s}).$$

The asymptotics are uniform for x in compacts of Ω.

Proof. Let us start with the easy case of $k > \frac{2s}{N-2s}$. In that case, since $H_a(x, \cdot)$ is uniformly bounded, we have

$$\lambda^{-\frac{k}{2}(N-2s)} \int_\Omega U_{x,\lambda}^{p-k} H_a(x, \cdot)^k \, dy \lesssim \lambda^{-k(N-2s)} \int_{B_{R\lambda}} U_{0,1}^{p-k} \, dy \lesssim \begin{cases} \lambda^{-k(N-2s)} & \text{if } k < \frac{N}{N-2s}, \\ \lambda^{-N} \ln \lambda & \text{if } k = \frac{N}{N-2s}, \\ \lambda^{-N} & \text{if } k > \frac{N}{N-2s}. \end{cases}$$

In any case, this is $o(\lambda^{-2s})$.

Now assume $1 \leq k \leq \frac{2s}{N-2s}$. Let us abbreviate $d = d(x)$ and $B_d = B_d(x)$ and show that the integral over $\Omega \setminus B_d$ is $o(\lambda^{-2s})$. Indeed, since $H_a(x, \cdot)$ is uniformly bounded,

$$\lambda^{-\frac{k}{2}(N-2s)} \int_{\Omega \setminus B_d} U_{x,\lambda}^{p-k} H_a(x, y)^k \, dy \lesssim \lambda^{-\frac{k}{2}(N-2s)} \int_{\mathbb{R}^N \setminus B_d} U_{x,\lambda}^{p-k} \, dy \qquad (A.10)$$

$$= \lambda^{-k(N-2s)} \int_{\mathbb{R}^N \setminus B_{d\lambda}} U_{0,1}^{p-k} \, dy = \lambda^{-N} = o(\lambda^{-2s}). \qquad (A.11)$$

To evaluate the remaining integral over B_d, we use the formula

$$(H_a(x, y))^k = \left(\phi_a(x) + \xi_x \cdot (y - x) - d_{N,s}(a(x) + o(1))|x - y|^{4s-N} \right)^k \qquad (A.12)$$

by Lemma A.3 (where ξ_x may be zero if we are in case (i) of that lemma). After multiplying out the right side, every term containing the factor $\xi_x \cdot (y - x)$ only once vanishes by oddness.

Let now $k \geq 2$. Since $\phi_a(x)$ and $a(x)$ are uniformly bounded and Ω is bounded, it is clear that we can estimate

$$H_a(x, y)^k = \phi_a(x)^k + \mathcal{O}(|y - x|^2 + |y - x|^{4s-N}) \leq \phi_a(x)^k + \mathcal{O}(|y - x|^{4s-N}).$$

For the last step we used that $4s - N \le 2 + 2s - N < 2$. Now

$$\lambda^{-\frac{k}{2}(N-2s)} \int_{B_d} U_{x,\lambda}^{p-k} |x-y|^{4s-N} \, \mathrm{d}y = \lambda^{-k(N-2s)} \lambda^{N-4s} \int_{B_{d\lambda}(0)} U_{0,1}^{p-k} |y|^{4s-N} \, \mathrm{d}y \tag{A.13}$$

$$\lesssim \lambda^{-N} \lambda^{-(k-2)(N-2s)} \times \begin{cases} \ln \lambda & \text{if } k = 2, \\ \lambda^{(k-2)(N-2s)} & \text{if } k > 2. \end{cases} \tag{A.14}$$

In any case, this is $o(\lambda^{-2s})$.

Finally, if $k = 1$, plugging in expansion (A.12), the term involving $a(x)$ is not negligible anymore. Instead, it gives

$$\lambda^{-\frac{N-2s}{2}} \int_{B_d} U_{x,\lambda}^{p-1} (a(x) + o(1)) |x-y|^{4s-N} \, \mathrm{d}y = \lambda^{-2s} a_{N,s} a(x) + o(\lambda^{-2s}),$$

which completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

APPENDIX B. AUXILIARY COMPUTATIONS

In this appendix, we collect some technical results and computations used throughout the paper.

First, we compute the L^q norm of $U_{x,\lambda}$ for various values of q.

Lemma B.1 (L^q-norm of $U_{x,\lambda}$). *Let $x \in \Omega$ and $q \in [1,\infty]$. As $\lambda \to \infty$, we have, uniformly for x in compact subsets,*

$$\|U_{x,\lambda}\|_{L^q(\Omega)} \sim \begin{cases} \lambda^{\frac{N-2s}{2} - \frac{N}{q}}, & q > \frac{N}{N-2s}, \\ \lambda^{-\frac{N-2s}{2}} (\ln \lambda)^{\frac{N-2s}{N}}, & q = \frac{N}{N-2s}, \\ \lambda^{-\frac{N-2s}{2}}, & q < \frac{N}{N-2s}. \end{cases}$$

Moreover, for $\partial_\lambda U_{x,\lambda} = \frac{N-2s}{2} \lambda^{\frac{N-2s-2}{2}} \frac{1 - \lambda^2 |x-y|^2}{(1+\lambda^2 |x-y|^2)^{\frac{N-2s+2}{2}}}$, we have $|\partial_\lambda U_{x,\lambda}| = \mathcal{O}(\lambda^{-1} U_{x,\lambda})$ pointwise and therefore

$$\|\partial_\lambda U_{x,\lambda}\|_q \lesssim \lambda^{-1} \|U_{x,\lambda}\|_q, \qquad q \in [1,\infty].$$

Finally, for $\partial_{x_i} U_{x,\lambda} = (-N+2s) \lambda^{\frac{N-2s+2}{2}} \frac{\lambda(x-y)}{(1+\lambda^2 |x-y|^2)^{\frac{N-2s+2}{2}}}$, we have

$$\|\partial_{x_i} U_{x,\lambda}\|_{L^q(\Omega)} \sim \begin{cases} \lambda^{\frac{N-2s+2}{2} - \frac{N}{q}}, & q > \frac{N}{N-2s+1}, \\ \lambda^{-\frac{N-2s}{2}} (\ln \lambda)^{\frac{N-2s+1}{N}}, & q = \frac{N}{N-2s+1}, \\ \lambda^{-\frac{N-2s}{2}}, & q < \frac{N}{N-2s+1}. \end{cases}$$

Lemma B.2. *We have*

$$\|(-\Delta)^{s/2} PU\| \sim 1, \qquad \|(-\Delta)^{s/2} \partial_\lambda PU\| \sim \lambda^{-1}, \qquad \|(-\Delta)^{s/2} \partial_{x_i} PU\| \sim \lambda.$$

Moreover,

$$\int_{\mathbb{R}^N} (-\Delta)^{s/2} PU_{x,\lambda} (-\Delta)^{s/2} \partial_\lambda PU_{x,\lambda} \, \mathrm{d}y \lesssim \lambda^{-N+2s-1},$$

$$\int_{\mathbb{R}^N} (-\Delta)^{s/2} PU_{x,\lambda} (-\Delta)^{s/2} \partial_{x_i} PU_{x,\lambda} \, \mathrm{d}y \lesssim \lambda^{-N+2s},$$

$$\int_{\mathbb{R}^N} (-\Delta)^{s/2} \partial_\lambda PU_{x,\lambda} (-\Delta)^{s/2} \partial_{x_i} PU_{x,\lambda} \, \mathrm{d}y \lesssim \lambda^{-N+2s-1},$$

$$\int_{\mathbb{R}^N} (-\Delta)^{s/2} \partial_{x_i} PU_{x,\lambda} (-\Delta)^{s/2} \partial_{x_j} PU_{x,\lambda} \, \mathrm{d}y \lesssim \lambda^{-N+2s}.$$

We remark that the bounds of Lemma B.2 are consistent with the ones proved in [34, Appendix B].

Lemma B.3. *We have*

$$0 \leq \frac{\lambda^{-\frac{N-2s}{2}}}{|x-y|^{N-2s}} - U_{x,\lambda}(y) = \lambda^{\frac{N-2s}{2}} h(\lambda(x-y)), \tag{B.1}$$

with

$$h(z) := \left(\frac{1}{1+|z|^2}\right)^{\frac{N-2s}{2}} - \frac{1}{|z|^{N-2s}}.$$

Moreover $h(z) \sim |z|^{-N-2+2s}$ *and* $|\nabla h(z)| \sim |z|^{-N+2s-3}$ *as* $|z| \to \infty$. *Consequently,* $h \in L^p(\mathbb{R}^N)$ *for every* $p \in [1, \frac{N}{N-2s})$ *and* $\nabla h \in L^p(\mathbb{R}^N)$ *for every* $p \in [1, \frac{N}{N-2s+1})$, *where the latter interval is possibly empty.*

Lemma B.4. *Let* $b \in C(\overline{\Omega}) \cap C^1(\Omega)$. *As* $\lambda \to \infty$, *uniformly for* x *in compact subsets of* Ω,

$$\int_\Omega b(y) U_{x,\lambda}(y) \lambda^{\frac{N-2s}{2}} h(\lambda(x-y)) \, \mathrm{d}y = \alpha_{N,s} \lambda^{-2s} b(x) + o(\lambda^{-2s}).$$

The numerical value of $\alpha_{N,s} = \int_{\mathbb{R}^N} U_{0,1}(y) h(y) \, \mathrm{d}y$ *is given in Lemma B.5 below.*

Proof. Abbreviate $d = d(x)$ and $B_d = B_d(x)$. We integrate separately over B_d and over $\Omega \setminus B_d$.

For the outer integral, from Lemma B.3 we get that $U_{0,1}(y) h(y) \sim |y|^{-2N+4s-2}$. Thus

$$\int_{\Omega \setminus B_d} b(y) U_{x,\lambda}(y) \lambda^{\frac{N-2s}{2}} h(\lambda(x-y)) \, \mathrm{d}y \lesssim \lambda^{-N+2s-2} = o(\lambda^{-N}) = o(\lambda^{-2s}).$$

For the inner integral, using that $b \in C^1(\Omega)$, we write $b(y) = b(x) + \nabla b(x) \cdot (y-x) + o(|x-y|)$ for $y \in B_d$. Then (the integral over $\nabla b(x) \cdot (y-x)$ cancels due to oddness)

$$\int_{B_d} b(y) U_{x,\lambda}(y) \lambda^{\frac{N-2s}{2}} h(\lambda(x-y)) \, \mathrm{d}y$$

$$= b(x) \lambda^{-2s} \int_{B_{\lambda d}(0)} U_{0,1}(y) h(y) \, \mathrm{d}y + o\left(\lambda^{-2s-1} \int_{B_{\lambda d}(0)} U_{0,1}(z) h(z) |z| \, \mathrm{d}z\right)$$

$$= b(x) \lambda^{-2s} \alpha_{N,s} + o(\lambda^{-2s} + o\left(\lambda^{-2s-1} \int_{B_{\lambda d}(0)} U_{0,1}(z) h(z) |z| \, \mathrm{d}z\right).$$

To show that the last term is $o(\lambda^{-2s})$ as well, note that by Lemma B.3 we have $U_{0,1}(z) h(z) |z| \lesssim |z|^{-2N+4s-1}$. Thus

$$\lambda^{-2s-1} \int_{B_{\lambda d}(0)} U_{0,1}(z) h(z) |z| \, \mathrm{d}z \lesssim \begin{cases} \lambda^{-2s-1} & \text{if } N > 4s-1, \\ \lambda^{-2s-1} \log \lambda & \text{if } N = 4s-1, \\ \lambda^{-N+4s-1} & \text{if } N < 4s-1. \end{cases}$$

This is $o(\lambda^{-2s})$ in all cases. $\qquad \square$

We compute explicitly the constants that appear in the asymptotic expansions throughout the paper.

Lemma B.5 (Constants). *For* $N > 2s$ *and* $p = \frac{2N}{N-2s}$, *let* $U_{0,1}(y) = \left(\frac{1}{1+|y|^2}\right)^{\frac{N-2s}{2}}$ *and* $h(y) = \frac{1}{|y|^{N-2s}} - \frac{1}{(1+|y|^2)^{\frac{N-2s}{2}}}$. *Then for every* $0 \leq k < \frac{N}{N-2s}$ *we have*

$$a_{N,s}(k) := \int_{\mathbb{R}^N} U_{0,1}(y)^{p-k} \, \mathrm{d}y = \frac{\pi^{N/2} \Gamma\left(\frac{N}{2}(1-k) + ks\right)}{\Gamma\left(\frac{N}{2}(2-k) + ks\right)}.$$

We denote $A_{N,s} := a_{N,s}(0)$ and $a_{N,s} := a_{N,s}(1)$. Further,

$$b_{N,s} := \int_{\mathbb{R}^N} U_{0,1}(y)^{\frac{N+2s}{N-2s}} |y|^{4s-N}\, \mathrm{d}y = \pi^{N/2} \frac{\Gamma(2s)\Gamma\left(\frac{N}{2}-s\right)}{\Gamma\left(\frac{N}{2}\right)\Gamma\left(\frac{N}{2}+s\right)},$$

$$\alpha_{N,s} := \int_{\mathbb{R}^N} U_{0,1}(y)h(y)\, \mathrm{d}y = \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}\Gamma\left(\frac{N}{2}-2s\right)\left(\frac{\Gamma(s)}{\Gamma(\frac{N}{2}-s)} - \frac{\Gamma(\frac{N}{2})}{\Gamma(N-2s)}\right),$$

Moreover, the constant in $(-\Delta)^s u(x) := C_{N,s} P.V. \int_{\mathbb{R}^N} \frac{u(x)-u(y)}{|x-y|^{N+2s}}\, \mathrm{d}y$ is given by

$$C_{N,s} := \frac{2^{2s}\Gamma(\frac{N+2s}{2})}{\pi^{N/2}s\Gamma(1-s)}$$

and the constant in $(-\Delta)^s U_{0,1} = c_{n,s}U_{0,1}^{\frac{N+2s}{N-2s}}$ is given by

$$c_{N,s} = 2^{2s}\frac{\Gamma(\frac{N+2s}{2})}{\Gamma(\frac{N-2s}{2})}.$$

The explicit value of the best fractional Sobolev constant in $\|(-\Delta)^{s/2}u\|^2 \geq S\|u\|^2_{\frac{2N}{N-2s}}$ is

$$S := S_{N,s} = 2^{2s}\pi^s \frac{\Gamma(\frac{N+2s}{2})}{\Gamma(\frac{N-2s}{2})}\left(\frac{\Gamma(N/2)}{\Gamma(N)}\right)^{2s/N}.$$

The constant in $(-\Delta)^s |x|^{4s-N} = -d_{N,s}^{-1}|x|^{2s-N}$ is given by

$$d_{N,s} := -2^{-2s}\frac{\Gamma(\frac{N-4s}{2})\Gamma(s)}{\Gamma(\frac{N-2s}{2})\Gamma(2s)} > 0. \tag{B.2}$$

The constant $\gamma_{N,s}$ in $((-\Delta)^s + a)G_a(x,\cdot) = \gamma_{N,s}\delta_x$ is given by

$$\gamma_{N,s} = \frac{2^{2s}\pi^{N/2}\Gamma(s)}{\Gamma(\frac{N-2s}{2})}.$$

Proof. The values of $a_{n,s}(k)$ and $b_{N,s}$ are a consequence of the following computation. For $\alpha, \beta > 0$,

$$\int_{\mathbb{R}^N}\left(\frac{1}{1+|y|^2}\right)^\alpha |y|^\beta\, \mathrm{d}y = \frac{2\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}\int_0^\infty \left(\frac{1}{1+r^2}\right)^\alpha r^{N-1+\beta}\, \mathrm{d}r$$

$$= \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}B\left(\frac{\beta+N}{2}, \alpha - \frac{\beta+N}{2}\right) = \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}\frac{\Gamma\left(\frac{\beta+N}{2}\right)\Gamma\left(\alpha - \frac{\beta+N}{2}\right)}{\Gamma(\alpha)}. \tag{B.3}$$

To compute $\alpha_{N,s}$, we write

$$\alpha_{N,s} = \int_{\mathbb{R}^N} U_{0,1}(y)h(y)\, \mathrm{d}y$$

$$= \frac{2\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}\int_0^\infty \underbrace{\left(\left(\frac{1}{1+r^2}\right)^{\frac{N-2s}{s}} r^{2s-1} - \left(\frac{1}{1+r^2}\right)^{N-2s} r^{N-1}\right)}_{I(r,N,s)}\, \mathrm{d}r.$$

If $N > 4s$, then the summands of $I(r, N, s)$ are separately integrable, in which case (B.3) gives

$$\alpha_{N,s} = \frac{\pi^{N/2}}{\Gamma\left(\frac{N}{2}\right)}\Gamma\left(\frac{N}{2}-2s\right)\left(\frac{\Gamma(s)}{\Gamma\left(\frac{N}{2}-s\right)} - \frac{\Gamma(\frac{N}{2})}{\Gamma(N-2s)}\right). \tag{B.4}$$

To extend this formula to the case $2s < N < 4s$ which concerns us, we remark that the right side of (B.4) defines a holomorphic function of s in the complex subdomain $\mathcal{D}_N := \{0 < \ \mathrm{Re}(s) < N/2\} \setminus$

$\{N/4\} \subset \mathbb{C}$. On the other hand, by a cancellation $I(r, N, s)$ remains integrable in $r \in (0, \infty)$ for every $s \in (0, 1)$ and $N \in (2s, 4s)$. Indeed,

$$I(r, N, s) \sim (r^{2s-1} - r^{N-1}) \qquad \text{as } r \to 0,$$

$$I(r, N, s) \sim \left(r^{-2\frac{N-2s}{2}} r^{2s-1} \left(\frac{1}{1 + \frac{1}{r^2}} \right)^{\frac{N-2s}{2}} - r^{-2(N-2s)} r^{N-1} \left(\frac{1}{1 + \frac{1}{r^2}} \right)^{N-2s} \right)$$

$$= \left(r^{-2\frac{N-2s}{2}} r^{2s-1} \left(1 + \frac{1}{r^2} \right)^{-\frac{N-2s}{2}} - r^{-2(N-2s)} r^{N-1} \left(1 + \frac{1}{r^2} \right)^{-(N-2s)} \right)$$

$$= r^{-N+4s-1} \left(1 - \frac{\frac{N}{2} - s}{r^2} + O(\frac{1}{r^4}) - 1 + \frac{N-2s}{r^2} + O(\frac{1}{r^4}) \right)$$

$$= \left(\frac{N}{2} - s \right) r^{-N+4s-3} + O(r^{-N+4s-5}) \qquad \text{as } r \to \infty.$$

By a standard argument, this implies that $\int_0^\infty I(r, N, s) \, dr$ is holomorphic in $\mathcal{D}_N$ as a function of s. By the identity theorem for analytic functions, the formula (B.4) thus holds also for $s \in (N/4, N/2)$, which is what we wanted to show.

Finally, the claimed value of S can be found e.g. in [16, Theorem 1.1] and that of $d_{N,s}$ in [30, Table 1, p. 168]. □

Lemma B.6. *Let* $2s < N < 4s$ *and let* $f_\varepsilon : (0, \infty) \to \mathbb{R}$ *be given by*

$$f_\varepsilon(\lambda) = \frac{A_\varepsilon}{\lambda^{2s}} - B_\varepsilon \frac{\varepsilon}{\lambda^{N-2s}}$$

with $A_\varepsilon, B_\varepsilon > 0$ *uniformly bounded away from 0 and* ∞. *The unique global minimum of* f_ε *is given by*

$$\lambda_0 = \lambda_0(\varepsilon) = \left(\frac{2sA_\varepsilon}{(N-2s)B_\varepsilon} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}}. \tag{B.5}$$

with corresponding minimal value

$$\min_{\lambda > 0} f_\varepsilon(\lambda) = f_\varepsilon(\lambda_0) = -\varepsilon^{\frac{2s}{4s-N}} \frac{B_\varepsilon^{\frac{2s}{4s-N}}}{A_\varepsilon^{\frac{N-2s}{4s-N}}} \left(\frac{N-2s}{2s} \right)^{\frac{2s}{4s-N}} \frac{4s-N}{N-2s}. \tag{B.6}$$

Moreover, there is a $c_0 > 0$ *such that, for all* $\varepsilon > 0$, *we have*

$$f_\varepsilon(\lambda) - f_\varepsilon(\lambda_0) \geq \begin{cases} c_0 \varepsilon^{\frac{2s-2}{4s-N}} \left(\lambda^{-1} - \lambda_0(\varepsilon)^{-1} \right)^2 & \text{if} \quad \left(\frac{A_\varepsilon}{B_\varepsilon} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}} \lambda^{-1} \leq 2 \left(\frac{2s}{N-2s} \right)^{\frac{1}{N-4s}}, \\ c_0 \varepsilon^{\frac{2s}{4s-N}} & \text{if} \quad \left(\frac{A_\varepsilon}{B_\varepsilon} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}} \lambda^{-1} > 2 \left(\frac{2s}{N-2s} \right)^{\frac{1}{N-4s}}. \end{cases} \tag{B.7}$$

Proof. The values of λ_0 and $f_\varepsilon(\lambda_0)$ are obtained by standard computations. Thus we only prove (B.7). Let $F(t) := t^{2s} - t^{N-2s}$ and denote by $t_0 := \left(\frac{2s}{N-2s} \right)^{-\frac{1}{4s-N}}$ the unique global minimum of F on $(0, \infty)$. Then, there exists $c > 0$ such that

$$F(t) - F(t_0) \geq \begin{cases} c(t - t_0)^2 & \text{if} \quad 0 < t \leq 2t_0, \\ ct_0^{N-2s} & \text{if} \quad t > 2t_0. \end{cases}$$

The assertion of the lemma now follows by rescaling. Indeed, it suffices to observe that

$$f_\varepsilon(\lambda) = A_\varepsilon^{-\frac{N-2s}{4s-N}} B_\varepsilon^{\frac{2s}{4s-N}} \varepsilon^{\frac{2s}{4s-N}} F\left(\left(\frac{A_\varepsilon}{B_\varepsilon} \right)^{\frac{1}{4s-N}} \varepsilon^{-\frac{1}{4s-N}} \lambda^{-1} \right)$$

and to use the boundedness of A_ε and B_ε. □

Acknowledgments

References

[1] N. Abatangelo. Large S-harmonic functions and boundary blow-up solutions for the fractional Laplacian. *Discrete Contin. Dyn. Syst.*, 35(12):5555–5607, 2015.

[2] W. Abdelhedi, H. Chtioui, and H. Hajaiej. The Bahri-Coron theorem for fractional Yamabe-type problems. *Adv. Nonlinear Stud.*, 18(2):393–407, 2018.

[3] A. Bahri and J.-M. Coron. On a nonlinear elliptic equation involving the critical Sobolev exponent: the effect of the topology of the domain. *Comm. Pure Appl. Math.*, 41(3):253–294, 1988.

[4] B. Barrios, E. Colorado, A. Pablo, and U. S'anchez. On some critical problems for the fractional Laplacian operator. *Journal of Differential Equations*, 252:6133–6162, 2011.

[5] W. Beckner. Sharp Sobolev inequalities on the sphere and the Moser-Trudinger inequality. *Ann. Math. (2)*, 138(1):213–242, 1993.

[6] C. Brändle, E. Colorado, A. de Pablo, and U. Sanchez. A concave—convex elliptic problem involving the fractional Laplacian. *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, 143:39 – 71, 2013.

[7] L. Brasco, D. Gómez-Castro, and J. L. Vázquez. Characterisation of homogeneous fractional Sobolev spaces. *Calc. Var. Partial Differ. Equ.*, 60(2):41, 2021. Id/No 60.

[8] H. Brezis. Elliptic equations with limiting Sobolev exponents—the impact of topology. *Communications on Pure and Applied Mathematics*, 39(S1):S17–S39, 1986.

[9] H. Brézis and L. Nirenberg. Positive solutions of nonlinear elliptic equations involving critical Sobolev exponents. *Comm. Pure Appl. Math.*, 36(4):437–477, 1983.

[10] H. Brezis and L. A. Peletier. Asymptotics for elliptic equations involving critical growth. In *Partial differential equations and the calculus of variations, Vol. I*, volume 1 of *Progr. Nonlinear Differential Equations Appl.*, pages 149–192. Birkhäuser Boston, Boston, MA, 1989.

[11] C. Bucur and E. Valdinoci. *Nonlocal diffusion and applications*, volume 20 of *Lecture Notes of the Unione Matematica Italiana*. Springer, [Cham]; Unione Matematica Italiana, Bologna, 2016.

[12] X. Cabré and J. Tan. Positive solutions of nonlinear problems involving the square root of the Laplacian. *Advances in Mathematics*, 224:2052–2093, 2009.

[13] L. Caffarelli and L. Silvestre. An extension problem related to the fractional Laplacian. *Communications in Partial Differential Equations*, 32:1245 – 1260, 2007.

[14] W. Choi and S. Kim. Classification of finite energy solutions to the fractional Lane-Emden-Fowler equations with slightly subcritical exponents. *Ann. Mat. Pura Appl. (4)*, 196(1):269–308, 2017.

[15] W. Choi, S. Kim, and K.-A. Lee. Asymptotic behavior of solutions for nonlinear elliptic problems with the fractional Laplacian. *J. Funct. Anal.*, 266(11):6531–6598, 2014.

[16] A. Cotsiolis and N. K. Tavoularis. Best constants for Sobolev inequalities for higher order fractional derivatives. *J. Math. Anal. Appl.*, 295(1):225–236, 2004.

[17] J. Dávila, L. López Ríos, and Y. Sire. Bubbling solutions for nonlocal elliptic problems. *Rev. Mat. Iberoam.*, 33(2):509–546, 2017.

[18] N. De Nitti and T. König. Energy asymptotics for the fractional Brézis–Nirenberg problem in high-dimension. *in preparation*, 2021.

[19] E. Di Nezza, G. Palatucci, and E. Valdinoci. Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.*, 136(5):521–573, 2012.

[20] O. Druet. Elliptic equations with critical Sobolev exponents in dimension 3. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 19(2):125–142, 2002.

[21] P. Esposito. On some conjectures proposed by Haïm Brezis. *Nonlinear Anal.*, 56(5):751–759, 2004.

[22] R. Frank, T. König, and H. Kovarik. Energy asymptotics in the Brezis-Nirenberg problem. The higher-dimensional case. *Mathematics in Engineering*, 2(1):119–140, 2020.

[23] R. L. Frank, T. König, and H. Kovařík. Blow-up of solutions of critical elliptic equation in three dimensions. *arXiv e-prints*, page arXiv:2102.10525, Feb. 2021.

[24] R. L. Frank, T. König, and H. Kovařík. Energy asymptotics in the three-dimensional Brezis-Nirenberg problem. *Calc. Var. Partial Differ. Equ.*, 60(2):47, 2021. Id/No 58.

[25] N. Garofalo. Fractional thoughts. In *New developments in the analysis of nonlocal operators*, volume 723 of *Contemp. Math.*, pages 1–135. Amer. Math. Soc., [Providence], RI, 2019.

[26] Y. Guo, B. Li, A. Pistoia, and S. Yan. The fractional Brezis-Nirenberg problems on lower dimensions. *J. Differential Equations*, 286:284–331, 2021.

[27] Z.-C. Han. Asymptotic approach to singular solutions for nonlinear elliptic equations involving critical Sobolev exponent. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 8(2):159–174, 1991.

[28] E. Hebey and M. Vaugon. From best constants to critical functions. *Math. Z.*, 237(4):737–767, 2001.

[29] E. Jannelli. The role played by space dimension in elliptic critical problems. *J. Differential Equations*, 156(2):407–426, 1999.

[30] M. Kwaśnicki. Fractional Laplace operator and its properties. In *Handbook of fractional calculus with applications. Vol. 1*, pages 159–193. De Gruyter, Berlin, 2019.

[31] E. H. Lieb and M. Loss. *Analysis*, volume 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2001.

[32] G. Palatucci and A. Pisante. Improved Sobolev embeddings, profile decomposition, and concentration-compactness for fractional Sobolev spaces. *Calc. Var. Partial Differential Equations*, 50(3-4):799–829, 2014.

[33] O. Rey. Proof of two conjectures of H. Brézis and L. A. Peletier. *Manuscripta Math.*, 65(1):19–37, 1989.

[34] O. Rey. The role of the Green's function in a nonlinear elliptic equation involving the critical Sobolev exponent. *J. Funct. Anal.*, 89(1):1–52, 1990.

[35] X. Ros-Oton and J. Serra. The extremal solution for the fractional Laplacian. *Calculus of variations and partial differential equations*, 50(3):723–750, 2014.

[36] R. Servadei and E. Valdinoci. A Brezis-Nirenberg result for non-local critical equations in low dimension. *Commun. Pure Appl. Anal.*, 12(6):2445–2464, 2013.

[37] R. Servadei and E. Valdinoci. On the spectrum of two different fractional operators. *Proc. Roy. Soc. Edinburgh Sect. A*, 144(4):831–855, 2014.

[38] R. Servadei and E. Valdinoci. The Brezis-Nirenberg result for the fractional Laplacian. *Trans. Amer. Math. Soc.*, 367(1):67–102, 2015.

[39] L. Silvestre. Regularity of the obstacle problem for a fractional power of the Laplace operator. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 60(1):67–112, 2007.

[40] P. Stinga and J. Torrea. Extension problem and Harnack's inequality for some fractional operators. *Communications in Partial Differential Equations*, 35:2092 – 2122, 2009.

[41] J. Tan. The Brezis–Nirenberg type problem involving the square root of the Laplacian. *Calculus of Variations and Partial Differential Equations*, 42:21–41, 2011.

(N. De Nitti) Friedrich-Alexander-Universität Erlangen-Nürnberg, Department of Data Science, Chair for Dynamics, Control and Numerics (Alexander von Humboldt Professorship), Cauerstr. 11, 91058 Erlangen, Germany.

Email address: nicola.de.nitti@fau.de

(T. König) Institut de Mathématiques de Jussieu – Paris Rive Gauche Université de Paris – Campus des Grands Moulins Bâtiment Sophie Germain, Boite Courrier 7012 8 Place Aurélie Nemours, 75205 Paris Cedex 13.

Email address: koenig@imj-prg.fr